BGP

Border Gateway Protocol (BGP)

Student Guide

Version 3.0

© 2003 KnowledgeNet.com, Inc. All Rights Reserved.

KNOWLEDGENET is a registered trademark; and the K DESIGN and THE BEST OF A NEW BREED are trademarks of KnowledgeNet.com, Inc. All other trademarks are the property of their respective companies.

MODULE 0 – COURSE INTRODUCTION **0-1**
Overview 0-1
Course Objectives 0-2
Cisco's Certification Track 0-4
Learner Skills and Knowledge 0-5
Learner Responsibilities 0-6
General Administration 0-7
Course Flow Diagram 0-8
Icons and Symbols 0-9
Learner Introductions 0-10

MODULE 1 – BGP OVERVIEW **1-1**
Overview 1-1
Outline 1-2

LESSON ONE: INTRODUCTION TO BGP **1-3**
Overview 1-3
Importance 1-3
Objectives 1-3
Learner Skills and Knowledge 1-4
Outline 1-4
Interdomain Routing 1-5
BGP Characteristics 1-9
Single-Homed Customers 1-15
Multihomed Customers 1-17
Transit Autonomous Systems 1-20
BGP Limitations 1-22
Summary 1-23
Next Steps 1-24
References 1-24
Assessment 1-25

LESSON TWO: BGP SESSION ESTABLISHMENT **1-29**
Overview 1-29
Importance 1-29
Objectives 1-29
Learner Skills and Knowledge 1-30
Outline 1-30
BGP Neighbor Discovery 1-31
Establishing a BGP Session 1-35
BGP Keepalives 1-39
MD5 Authentication 1-42
Summary 1-43
Next Steps 1-43
References 1-43
Assessment 1-44

LESSON THREE: BGP ROUTE PROCESSING **1-47**
Overview 1-47
Importance 1-47
Objectives 1-47
Learner Skills and Knowledge 1-48

Copyright © 2003, Cisco Systems, Inc.

Outline 1-48
Receiving Routing Updates 1-49
Building the BGP Table 1-51
BGP Route Selection Criteria 1-53
BGP Route Propagation 1-56
Building the IP Routing Table 1-58
Advertising Local Networks 1-60
Automatic Summarization 1-63
Summary 1-67
Next Steps 1-67
References 1-67
Assessment 1-68

LESSON FOUR: BASIC BGP CONFIGURATION 1-71
Overview 1-71
Importance 1-71
Objectives 1-71
Learner Skills and Knowledge 1-72
Outline 1-72
BGP Routing Process 1-73
Configuring External Neighbors 1-75
Configuring BGP Timers 1-78
Configuring MD5 Authentication 1-80
Announcing Networks in BGP 1-82
Redistributing Routes into BGP 1-87
Configuring Classless BGP 1-93
Aggregating BGP Networks 1-98
Multihomed Customer Problem 1-105
Summary 1-108
Next Steps 1-109
References 1-109
Assessment 1-110

LESSON FIVE: MONITORING AND TROUBLESHOOTING BGP 1-113
Overview 1-113
Importance 1-113
Objectives 1-113
Learner Skills and Knowledge 1-114
Outline 1-114
Monitoring Overall BGP Routing 1-115
Monitoring BGP Neighbors 1-117
Monitoring the BGP Table 1-119
Debugging BGP 1-122
BGP Session Startup Problems 1-126
BGP Neighbor Not Reachable 1-127
BGP Neighbor Not Configured 1-130
BGP AS Number Mismatch 1-132
Summary 1-134
Next Steps 1-135
References 1-135
Assessment 1-136

MODULE 2 – ROUTE SELECTION USING POLICY CONTROLS 2-1
Overview 2-1

 Copyright © 2003, Cisco Systems, Inc.

Outline 2-2

LESSON ONE: MULTIHOMED BGP 2-3

Overview 2-3
Importance 2-3
Objectives 2-3
Learner Skills and Knowledge 2-4
Outline 2-4
Business Requirements of Multihomed BGP Networks 2-5
Technical Requirements of Multihomed BGP Networks 2-7
BGP Route Selection Without Policies 2-9
Multihomed Customer Routing Policies 2-11
Influencing BGP Route Selection 2-13
BGP Filters Transit Traffic Issue 2-15
Summary 2-19
Next Steps 2-20
References 2-20
Assessment 2-21

LESSON TWO: AS-PATH FILTERS 2-27

Overview 2-27
Importance 2-27
Objectives 2-27
Learner Skills and Knowledge 2-28
Outline 2-28
AS-Path Filtering Scenarios 2-29
AS-Path Regular Expressions 2-31
String Matching Regular Expressions 2-33
Applying AS-Path Filters 2-43
Configuring BGP AS-Path Filters 2-44
Monitoring AS-Path Filters 2-47
Summary 2-53
Next Steps 2-54
References 2-54
Assessment 2-55

LESSON THREE: PREFIX-LIST FILTERS 2-61

Overview 2-61
Importance 2-61
Objectives 2-61
Learner Skills and Knowledge 2-62
Outline 2-62
Requirements for Prefix-Based Filters 2-63
Prefix-Lists vs. IP Access-Lists 2-65
Configuring Prefix-Lists 2-68
BGP Filters Implementation 2-74
Implementing Prefix-Lists in the BGP Process 2-75
Modifying Prefix-Lists 2-80
Monitoring Prefix-Lists 2-82
Summary 2-86
Next Steps 2-87
References 2-87
Assessment 2-88

Copyright © 2003, Cisco Systems, Inc.

LESSON FOUR: OUTBOUND ROUTE FILTERING 2-95
Overview 2-95
Importance 2-95
Objectives 2-96
Learner Skills and Knowledge 2-96
Outline 2-96
Outbound Route Filter 2-97
Outbound Route Filter Message 2-99
Configuring Outbound Route Filtering 2-103
Using Outbound Route Filtering 2-106
Monitoring Outbound Route Filtering 2-108
Summary 2-109
Next Steps 2-109
References 2-109
Assessment 2-110

LESSON FIVE: ROUTE-MAPS AS BGP FILTERS 2-113
Overview 2-113
Importance 2-113
Objectives 2-113
Learner Skills and Knowledge 2-114
Outline 2-114
Why Use Route-Maps as BGP Filters? 2-115
Route-Maps Overview 2-116
Prefix-List Use in Route-Maps 2-121
BGP Filters 2-123
Using Route-Maps as BGP Filters 2-124
Monitoring Route-Maps 2-125
Summary 2-131
Next Steps 2-132
References 2-132
Assessment 2-133

LESSON SIX: IMPLEMENTING CHANGES IN BGP POLICY 2-137
Overview 2-137
Importance 2-137
Objectives 2-138
Learner Skills and Knowledge 2-138
Outline 2-138
Traditional Filtering Limitations 2-139
BGP Soft Reconfiguration 2-141
Cisco IOS Commands for Soft Reconfiguration 2-144
Monitoring Soft Reconfiguration 2-148
Route Refresh 2-150
Using Route Refresh 2-154
Monitoring Route Refresh 2-156
Summary 2-159
Next Steps 2-160
References 2-160
Assessment 2-161

MODULE 3 – ROUTE SELECTION USING ATTRIBUTES 3-1
Overview 3-1
Outline 3-2

 Copyright © 2003, Cisco Systems, Inc.

LESSON ONE: BGP PATH ATTRIBUTES 3-3

Overview 3-3
Importance 3-3
Objectives 3-3
Learner Skills and Knowledge 3-4
Outline 3-4
BGP Path Attributes 3-5
Well-Known BGP Attributes 3-6
Optional BGP Attributes 3-10
AS-Path Attribute 3-13
Next-Hop Attribute 3-16
Summary 3-21
Next Steps 3-21
References 3-21
Assessment 3-22

LESSON TWO: CATALYST 3550 INTERFACE CONFIGURATION 3-25

Overview 3-25
Importance 3-25
Objectives 3-25
Learner Skills and Knowledge 3-26
Outline 3-26
BGP Route Selection Criteria 3-27
Influencing BGP Route Selection 3-28
Configuring Per-Neighbor Weights 3-30
Changing Weights with AS-Path Filters 3-35
Changing Weights with Route-Maps 3-40
Monitoring BGP Route Selection and Weights 3-43
BGP Route Selection and Filtering Tools Summary 3-46
Summary 3-47
Next Steps 3-48
References 3-48
Assessment 3-49

LESSON THREE: BGP LOCAL PREFERENCE 3-55

Overview 3-55
Importance 3-55
Objectives 3-55
Learner Skills and Knowledge 3-56
Outline 3-56
Consistent Route Selection Within the AS 3-57
BGP Local Preference 3-61
Configuring Default Local Preference 3-64
Configuring Local Preference with Route-Maps 3-66
Monitoring Local Preference 3-68
Summary 3-73
Next Steps 3-73
References 3-73
Assessment 3-74

LESSON FOUR: AS-PATH PREPENDING 3-79

Overview 3-79

Copyright © 2003, Cisco Systems, Inc.

Importance 3-79
Objectives 3-79
Learner Skills and Knowledge 3-80
Outline 3-80
Return Path Selection in a Multihomed AS 3-81
AS-Path Prepending 3-86
AS-Path Prepending Design Considerations 3-90
Configuring AS-Path Prepending 3-92
Monitoring AS-Path Prepending 3-96
AS-Path Filtering Concerns with AS-Path Prepending 3-98
Summary 3-101
Next Steps 3-102
References 3-102
Assessment 3-103

LESSON FIVE: BGP MULTI-EXIT DISCRIMINATOR 3-107

Overview 3-107
Importance 3-107
Objectives 3-107
Learner Skills and Knowledge 3-108
Outline 3-108
Selecting the Proper Return Path 3-109
MED Propagation in a BGP Network 3-112
Changing Default MED 3-113
Changing MED with Route-Maps 3-115
Advanced MED Configuration 3-117
Monitoring MED 3-121
Troubleshooting MED 3-125
Summary 3-129
Next Steps 3-130
References 3-130
Assessment 3-131

LESSON SIX: BGP COMMUNITIES 3-137

Overview 3-137
Importance 3-137
Objectives 3-138
Learner Skills and Knowledge 3-138
Outline 3-139
Selecting the Proper Return Path 3-140
BGP Communities Overview 3-142
Using Communities 3-145
Configuring BGP Communities 3-148
Configuring Route Tagging with BGP Communities 3-150
Configuring Community Propagation 3-153
Defining BGP Community-Lists 3-155
Matching BGP Communities in Route-Maps 3-158
Monitoring Communities 3-161
Summary 3-165
Next Steps 3-166
References 3-166
Assessment 3-167

 Copyright © 2003, Cisco Systems, Inc.

MODULE 4 – CUSTOMER-TO-PROVIDER CONNECTIVITY WITH BGP 4-1

Overview 4-1
Outline 4-2

LESSON ONE: CUSTOMER-TO-PROVIDER CONNECTIVITY REQUIREMENTS 4-3

Overview 4-3
Importance 4-3
Objectives 4-3
Learner Skills and Knowledge 4-4
Outline 4-4
Customer Connectivity Types 4-5
Redundancy in Customer Connections 4-7
Customer-to-Provider Routing Schemes 4-12
Customer Routing Single - Permanent Connection 4-14
Addressing Requirements - Single-Homed Customers 4-19
AS Number Allocation - Single-Homed Customers 4-23
Summary 4-26
Next Steps 4-27
References 4-27
Assessment 4-28

LESSON TWO: IMPLEMENTING CUSTOMER CONNECTIVITY USING STATIC ROUTING 4-31

Overview 4-31
Importance 4-31
Objectives 4-32
Learner Skills and Knowledge 4-32
Outline 4-33
Why Use Static Routing? 4-34
Characteristics of Static Routing 4-36
Designing Static Route Propagation in a Service Provider Network 4-40
Static Route Propagation Case Study Parameters 4-42
BGP Backup with Static Routes 4-47
Floating Static Routes with BGP 4-53
Load Sharing with Static Routes 4-58
Summary 4-62
Next Steps 4-63
References 4-63
Assessment 4-64

LESSON THREE: CONNECTING A MULTIHOMED CUSTOMER TO A SINGLE SERVICE PROVIDER 4-69

Overview 4-69
Importance 4-69
Objectives 4-70
Learner Skills and Knowledge 4-70
Outline 4-71
Configuring BGP on Multihomed Customer Routers 4-72
Conditional Advertising in Multihomed Customer Networks 4-76
Configuring BGP on Service Provider Routers 4-79
Removing Private AS Numbers 4-83

Copyright © 2003, Cisco Systems, Inc.

Backup Solutions with BGP 4-87
Load Sharing with the Multihomed Customer 4-90
Load Sharing with BGP Multipath 4-92
Load Sharing with EBGP Multihop 4-94
Summary 4-98
Next Steps 4-99
References 4-99
Assessment 4-100

LESSON FOUR: CONNECTING A MULTIHOMED CUSTOMER TO MULTIPLE SERVICE PROVIDERS 4-105

Overview 4-105
Importance 4-105
Objectives 4-106
Learner Skills and Knowledge 4-106
Outline 4-107
Configuring BGP for Multihomed Customers 4-108
Multihomed Customer Address Space Selection 4-111
Multihomed Customer AS Number Selection 4-112
AS Number Translation 4-114
Primary/Backup Link Selection 4-116
BGP Incoming Link Selection 4-118
Load Sharing with Multiple Providers 4-123
Summary 4-125
Next Steps 4-126
References 4-126
Assessment 4-127

MODULE 5 – BGP TRANSIT AUTONOMOUS SYSTEMS 5-1

Overview 5-1
Outline 5-2

LESSON ONE: TRANSIT AUTONOMOUS SYSTEM FUNCTIONS 5-3

Overview 5-3
Importance 5-3
Objectives 5-4
Learner Skills and Knowledge 5-4
Outline 5-4
Transit Autonomous System Tasks 5-5
External Route Propagation 5-6
Internal Route Propagation 5-8
Packet Forwarding in an Autonomous System 5-9
Core Router IBGP Requirements in a Transit Autonomous 5-11
Summary 5-13
Next Steps 5-13
References 5-13
Assessment 5-14

LESSON TWO: IBGP AND EBGP INTERACTION IN A TRANSIT AUTONOMOUS SYSTEM 5-17

Overview 5-17
Importance 5-17
Objectives 5-17
Learner Skills and Knowledge 5-18

 Copyright © 2003, Cisco Systems, Inc.

Outline 5-18
AS-Path Processing in IBGP 5-19
BGP Split Horizon 5-21
IBGP Full Mesh 5-23
IBGP Neighbors 5-26
IBGP Next-Hop Processing 5-29
Transit Network Using External Next Hops 5-31
Transit Network Using Edge Routers as Next Hops 5-33
Differences Between EBGP and IBGP Sessions 5-37
Summary 5-41
Next Steps 5-42
References 5-42
Assessment 5-43

LESSON THREE: PACKET FORWARDING IN TRANSIT AUTONOMOUS SYSTEMS 5-47

Overview 5-47
Importance 5-47
Objectives 5-47
Learner Skills and Knowledge 5-48
Outline 5-48
Packet Forwarding in a Transit Autonomous System 5-49
Recursive Lookup in Cisco IOS Software 5-53
Routing Protocols in a Transit Autonomous System 5-55
BGP and IGP Interaction 5-58
Problems with BGP and IGP Interaction 5-61
Summary 5-62
Next Steps 5-62
References 5-62
Assessment 5-63

LESSON FOUR: CONFIGURING A TRANSIT AUTONOMOUS SYSTEM 5-67

Overview 5-67
Importance 5-67
Objectives 5-67
Learner Skills and Knowledge 5-68
Outline 5-68
Configuring IBGP Neighbors 5-69
Configuring IBGP Sessions Between Loopback Interfaces 5-71
Configuring BGP Synchronization 5-73
Changing the Administrative Distance of BGP Routes 5-75
Scalability Limitations of IBGP-Based Transit Backbones 5-77
Summary 5-79
Next Steps 5-79
References 5-79
Assessment 5-80

LESSON FIVE: MONITORING AND TROUBLESHOOTING IBGP IN TRANSIT AUTONOMOUS SYSTEMS 5-83

Overview 5-83
Importance 5-83
Objectives 5-83

Copyright © 2003, Cisco Systems, Inc.

Learner Skills and Knowledge 5-84
Outline 5-84
Monitoring IBGP 5-85
Common IBGP Problems 5-89
Troubleshooting IBGP Session Startup Issues 5-90
Troubleshooting IBGP Route Selection Issues 5-93
Troubleshooting IBGP Synchronization Issues 5-95
Summary 5-96
Next Steps 5-96
References 5-96
Assessment 5-97

MODULE 6 – SCALING SERVICE PROVIDER NETWORKS 6-1
Overview 6-1
Outline 6-2

LESSON ONE: SCALING IGP AND BGP IN SERVICE PROVIDER NETWORKS 6-3
Overview 6-3
Importance 6-3
Objectives 6-4
Learner Skills and Knowledge 6-4
Outline 6-4
Common Service Provider Network 6-5
Route Propagation in Service Provider Networks 6-10
Scaling Service Provider Routing Protocols 6-15
Scaling Service Provider Addressing 6-19
Summary 6-22
Next Steps 6-22
References 6-22
Assessment 6-23

LESSON TWO: INTRODUCTION TO ROUTE REFLECTORS 6-25
Overview 6-25
Importance 6-25
Objectives 6-25
Learner Skills and Knowledge 6-26
Outline 6-26
IBGP Scalability Issues in a Transit AS 6-27
Route Reflector Split-Horizon Rules 6-29
Redundant Route Reflectors 6-33
Route Reflector Clusters 6-35
Additional Route Reflector Loop Prevention Mechanisms 6-38
Summary 6-40
Next Steps 6-40
References 6-40
Assessment 6-41

LESSON THREE: NETWORK DESIGN WITH ROUTE REFLECTORS 6-45
Overview 6-45
Importance 6-53
Objectives 6-45
Learner Skills and Knowledge 6-46
Outline 6-46

 Copyright © 2003, Cisco Systems, Inc.

Network Design with Route Reflectors 6-47
Potential Network Issues 6-51
Hierarchical Route Reflectors 6-53
Summary 6-55
Next Steps 6-55
References 6-55
Assessment 6-56

LESSON FOUR: CONFIGURING AND MONITORING ROUTE REFLECTORS 6-59

Overview 6-59
Importance 6-59
Objectives 6-59
Learner Skills and Knowledge 6-60
Outline 6-60
Route Reflector Backbone Migration 6-61
Configuring Route Reflectors 6-63
Monitoring Route Reflectors 6-67
Summary 6-71
Next Steps 6-71
References 6-71
Assessment 6-72

LESSON FIVE: INTRODUCTION TO CONFEDERATIONS 6-77

Overview 6-77
Importance 6-77
Objectives 6-77
Learner Skills and Knowledge 6-78
Outline 6-78
IBGP Transit AS Problems 6-79
Splitting a Transit AS with BGP Confederations 6-81
AS-Path Propagation Within the BGP Confederation 6-83
AS-Path Processing in BGP Confederations 6-86
Intraconfederation EBGP Session Properties 6-88
Summary 6-90
Next Steps 6-90
References 6-90
Assessment 6-91

LESSON SIX: CONFIGURING AND MONITORING CONFEDERATIONS 6-95

Overview 6-95
Importance 6-95
Objectives 6-95
Learner Skills and Knowledge 6-96
Outline 6-96
BGP Confederation Design Rules 6-97
Planning BGP Confederations 6-99
Configuring BGP Confederations 6-101
Monitoring BGP Confederations 6-107
Summary 6-112
Next Steps 6-112
References 6-112
Assessment 6-113

Copyright © 2003, Cisco Systems, Inc.

MODULE 7 – OPTIMIZING BGP SCALABILITY 7-1
Overview 7-1
Outline 7-2

LESSON ONE: IMPROVING BGP CONVERGENCE 7-3
Overview 7-3
Importance 7-3
Objectives 7-4
Learner Skills and Knowledge 7-4
Outline 7-5
BGP Convergence 7-6
BGP Processes 7-8
CPU Effects of BGP Processes 7-10
Improving BGP Convergence 7-13
Path MTU Discovery 7-17
Increasing Input Queue Depth 7-20
BGP Scan Time 7-24
BGP Advertisement Interval 7-27
Summary 7-30
Next Steps 7-30
References 7-30
Assessment 7-31

LESSON TWO: LIMITING THE NUMBER OF PREFIXES RECEIVED FROM A BGP NEIGHBOR 7-37
Overview 7-37
Importance 7-37
Objectives 7-37
Learner Skills and Knowledge 7-38
Outline 7-38
Limiting the Number of Routes Received from a Neighbor 7-39
Configuring the BGP Maximum-Prefix Function 7-41
Monitoring the BGP Maximum-Prefix Function 7-43
Summary 7-47
Next Steps 7-47
References 7-47
Assessment 7-48

LESSON THREE: BGP PEER GROUPS 7-55
Overview 7-55
Importance 7-55
Objectives 7-55
Learner Skills and Knowledge 7-56
Outline 7-56
Peer Group Requirements 7-57
Peer Groups as a BGP Performance Tool 7-62
BGP Peer Group Limitations 7-64
BGP Peer Groups in Cisco IOS Software 7-66
Configuring Peer Groups 7-68
Monitoring Peer Groups 7-75
Summary 7-80
Next Steps 7-81
References 7-81

 Copyright © 2003, Cisco Systems, Inc.

Assessment 7-82

LESSON FOUR: BGP ROUTE DAMPENING 7-87

Overview 7-87
Importance 7-87
Objectives 7-88
Learner Skills and Knowledge 7-88
Outline 7-88
BGP Route Dampening 7-89
Route-Dampening Operation 7-91
Configuring BGP Route Dampening 7-94
Releasing Dampened Routes 7-99
Monitoring Route Dampening 7-101
Summary 7-110
Next Steps 7-110
References 7-110
Assessment 7-111

ADDITIONAL RESOURCES

Appendix A: Answer Key **A-1**
Appendix B: Course Glossary **B-1**

Course Introduction

Overview

The *Configuring BGP on Cisco Routers* (BGP) v3.0 course provides students with in-depth knowledge of Border Gateway Protocol (BGP), the routing protocol that is one of the underlying foundations of the Internet and New World technologies such as Multiprotocol Layer Switching (MPLS). This curriculum covers the theory of BGP, configuration of BGP on Cisco IOS® routers, detailed troubleshooting information, and hands-on exercises that provide learners with the skills needed to configure and troubleshoot BGP networks in customer environments. Different service solutions in the curriculum cover BGP network design issues and usage rules for various BGP features, preparing learners to design and implement efficient, optimal, and trouble-free BGP networks.

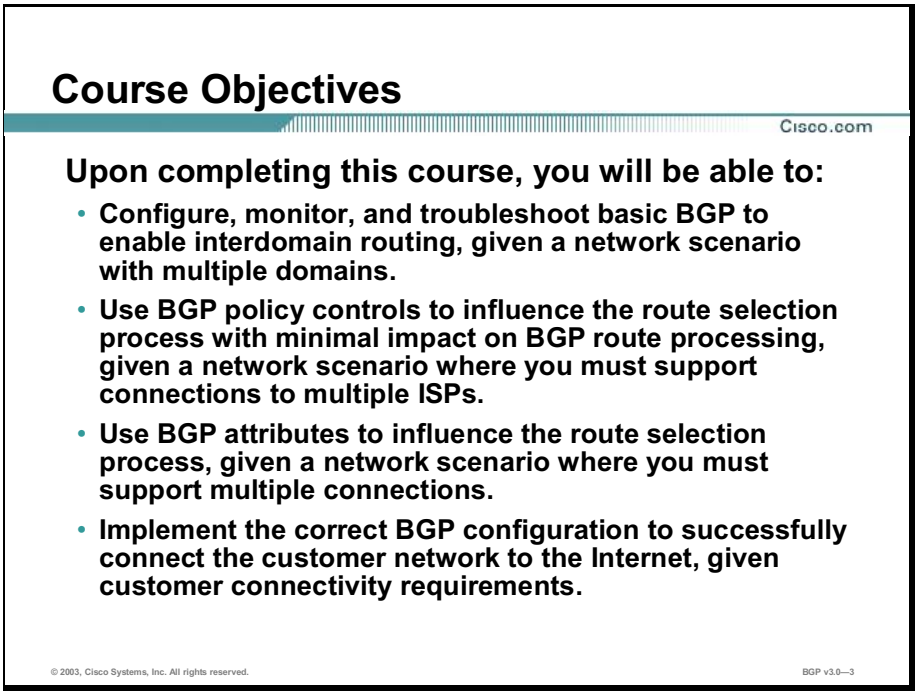
Outline

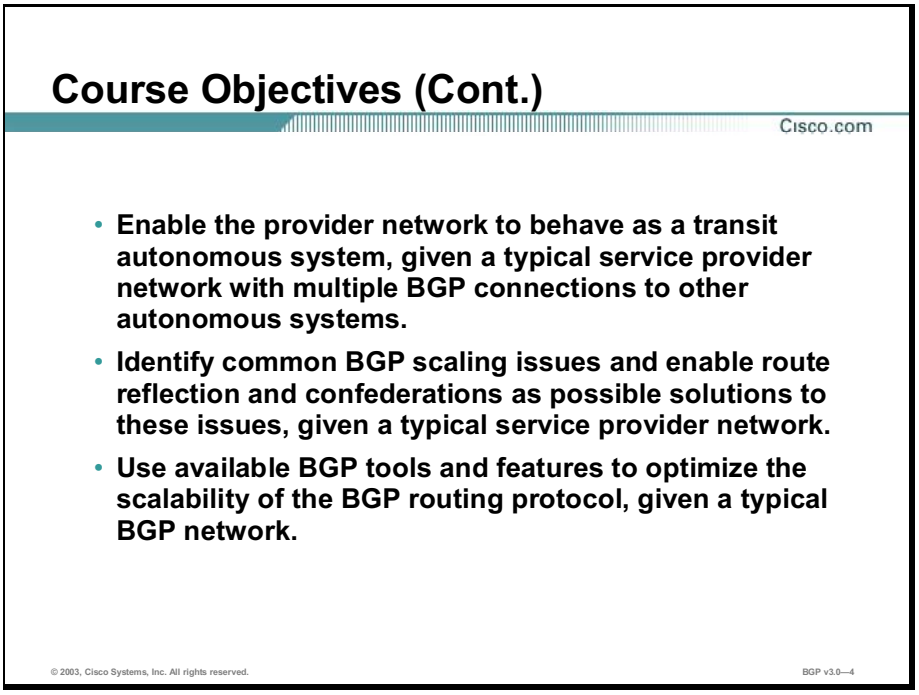
The Course Introduction includes these topics:

- Course Objectives
- Cisco Certifications
- Learner Skills and Knowledge
- Learner Responsibilities
- General Administration
- Course Flow Diagram
- Icons and Symbols
- Learner Introductions

Course Objectives

This topic lists the course objectives.


Course Objectives
Cisco.com
Upon completing this course, you will be able to:
• Configure, monitor, and troubleshoot basic BGP to enable interdomain routing, given a network scenario with multiple domains.
• Use BGP policy controls to influence the route selection process with minimal impact on BGP route processing, given a network scenario where you must support connections to multiple ISPs.
• Use BGP attributes to influence the route selection process, given a network scenario where you must support multiple connections.
• Implement the correct BGP configuration to successfully connect the customer network to the Internet, given customer connectivity requirements.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3


Course Objectives (Cont.)
Cisco.com
• Enable the provider network to behave as a transit autonomous system, given a typical service provider network with multiple BGP connections to other autonomous systems.
• Identify common BGP scaling issues and enable route reflection and confederations as possible solutions to these issues, given a typical service provider network.
• Use available BGP tools and features to optimize the scalability of the BGP routing protocol, given a typical BGP network.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4

 Copyright © 2003, Cisco Systems, Inc.

Upon completing this course, you will be able to:

- Configure, monitor, and troubleshoot basic BGP to enable interdomain routing, given a network scenario with multiple domains
- Use BGP policy controls to influence the route selection process with minimal impact on BGP route processing, given a network scenario where you must support connections to multiple ISPs
- Use BGP attributes to influence the route selection process, given a network scenario where you must support multiple connections
- Implement the correct BGP configuration to successfully connect the customer network to the Internet, given a network scenario where you must support multiple connections
- Enable the provider network to behave as a transit autonomous system, given a typical service provider network with multiple BGP connections to other autonomous systems
- Identify common BGP scaling issues and enable route reflection and confederations as possible solutions to these issues, given a typical service provider network with multiple BGP connections to other autonomous systems
- Use available BGP tools and features to optimize the scalability of the BGP routing protocol, given a typical BGP network

Copyright © 2003, Cisco Systems, Inc.

Cisco Certifications

This topic lists the certification requirements of this course.

This education offering is part of the Cisco Certified Internetwork Professional (CCIP) program.

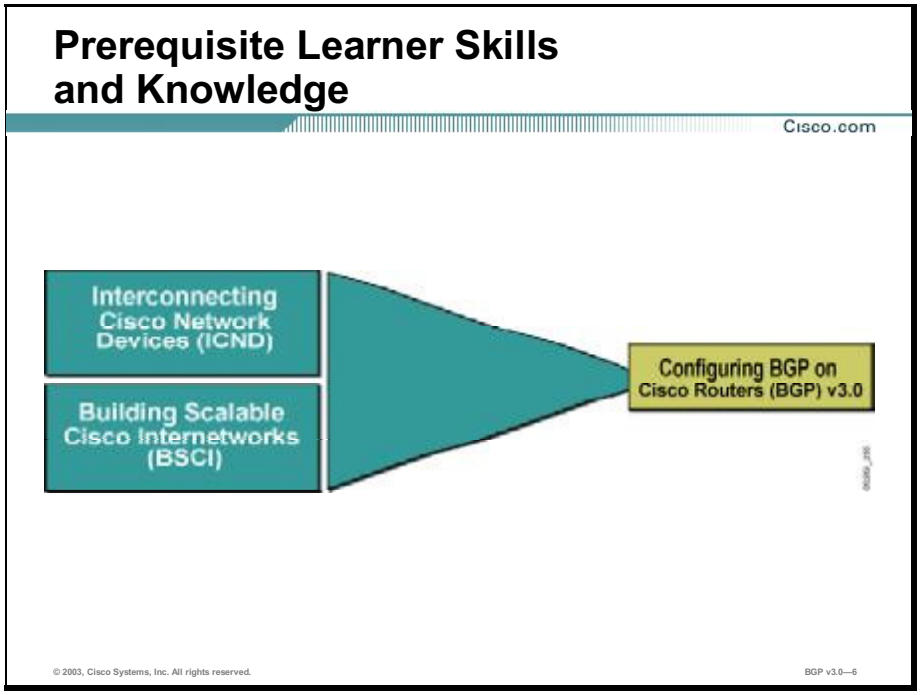
- The CCIP certification indicates advanced or journeyman knowledge of the technologies and solutions for providing data/voice communication and services.
- The CCIP certification provides individuals with competencies in infrastructure or access solutions in a Cisco end-to-end environment.
- CCIP professionals have a detailed understanding of Internet Protocol technologies represented by the Cisco courses *Building Scalable Cisco Internetworks*, *Implementing Cisco Multicast*, and *Implementing Cisco QoS*, as well as one or more specialized areas, including MPLS, Metro (Optical), Packet Telephony, DSL, Cable, Content Networking, and Security.

There is no prerequisite certification for CCIP; however, a valid Cisco Certified Network Associate (CCNA) certification is highly recommended.

 Copyright © 2003, Cisco Systems, Inc.

Learner Skills and Knowledge

This topic lists the course prerequisites.

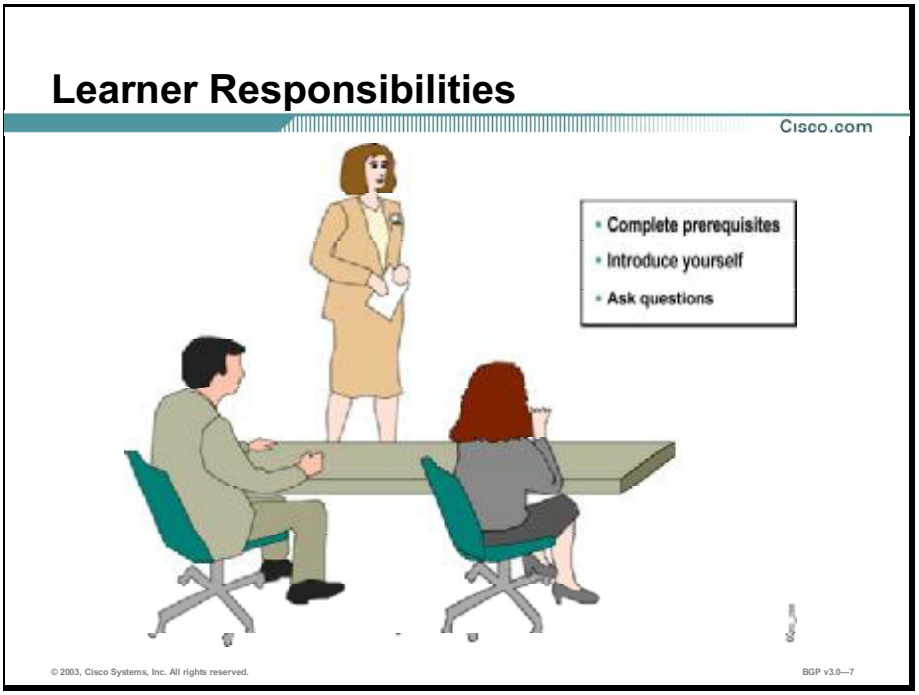

Prerequisite Learner Skills and Knowledge
Cisco.com
Interconnecting Cisco Network Devices (ICND)
Building Scalable Cisco Internetworks (BSCI)
Configuring BGP on Cisco Routers (BGP) v3.0
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6

To fully benefit from this course, you must have these prerequisite skills and knowledge:

- Completion of the *Interconnecting Cisco Networking Devices* (ICND) course or CCNA certification
- Completion of the *Building Scalable Cisco Internetworks* (BSCI) course or equivalent

Learner Responsibilities

This topic discusses the responsibilities of the learners.



To take full advantage of the information presented in this course, you must have completed the prerequisite requirements.

In class, you are expected to participate in all lesson exercises and assessments.

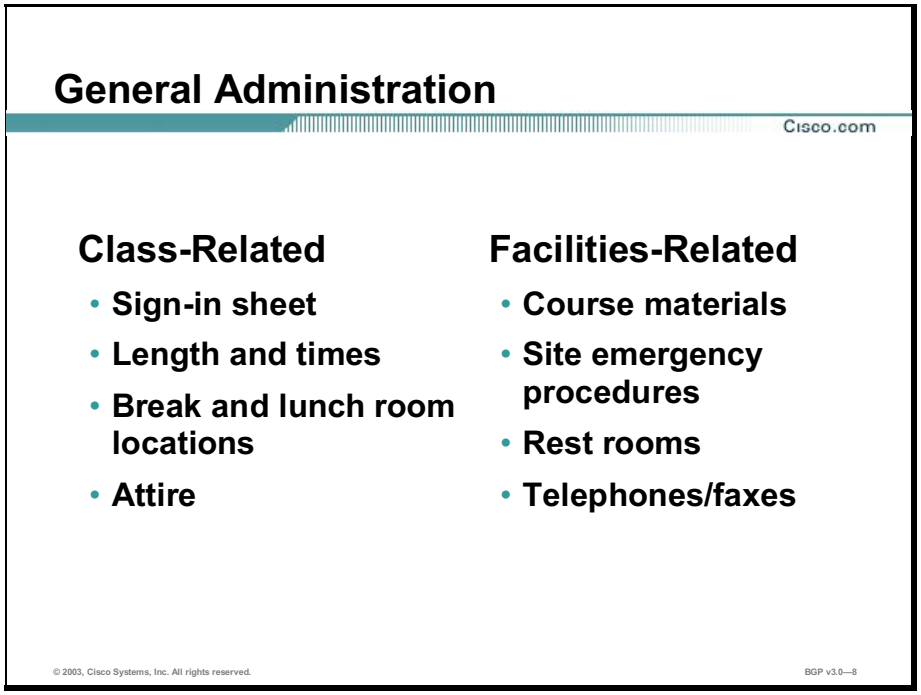
In addition, you are encouraged to ask any questions relevant to the course materials.

If you have pertinent information or questions concerning future Cisco product releases and product features, please discuss these topics during breaks or after class. The instructor will answer your questions or direct you to an appropriate information source.

 Copyright © 2003, Cisco Systems, Inc.

General Administration

This topic lists the administrative issues for the course.

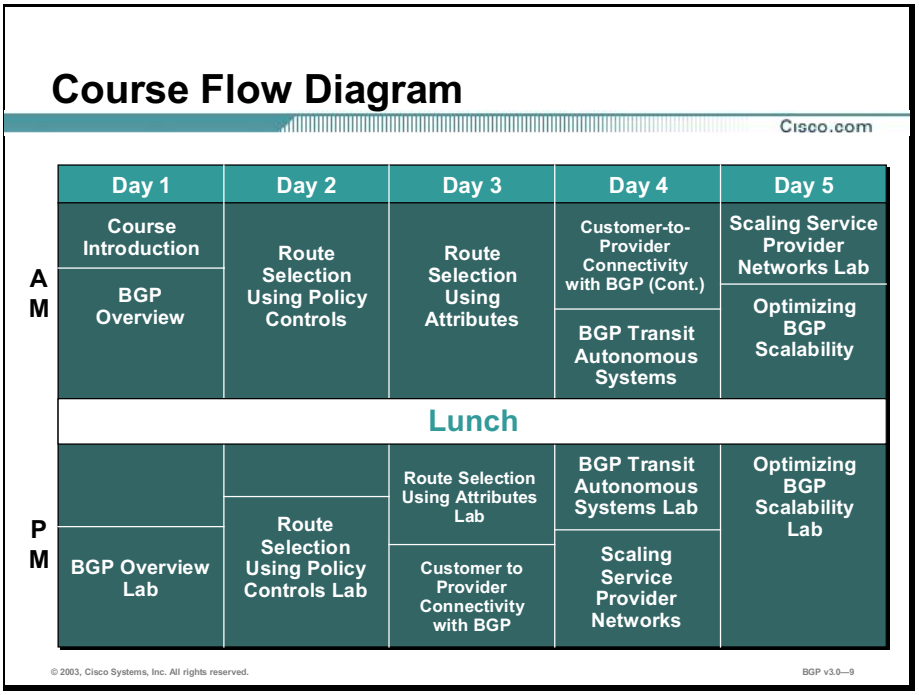


The instructor will discuss the administrative issues noted here so you know exactly what to expect from the class.

- Sign-in process
- Starting and anticipated ending times of each class day
- Class breaks and lunch facilities
- Appropriate attire during class
- Materials you can expect to receive during class
- What to do in the event of an emergency
- Location of the rest rooms
- How to send and receive telephone and fax messages

Course Flow Diagram

This topic covers the suggested flow of the course materials.

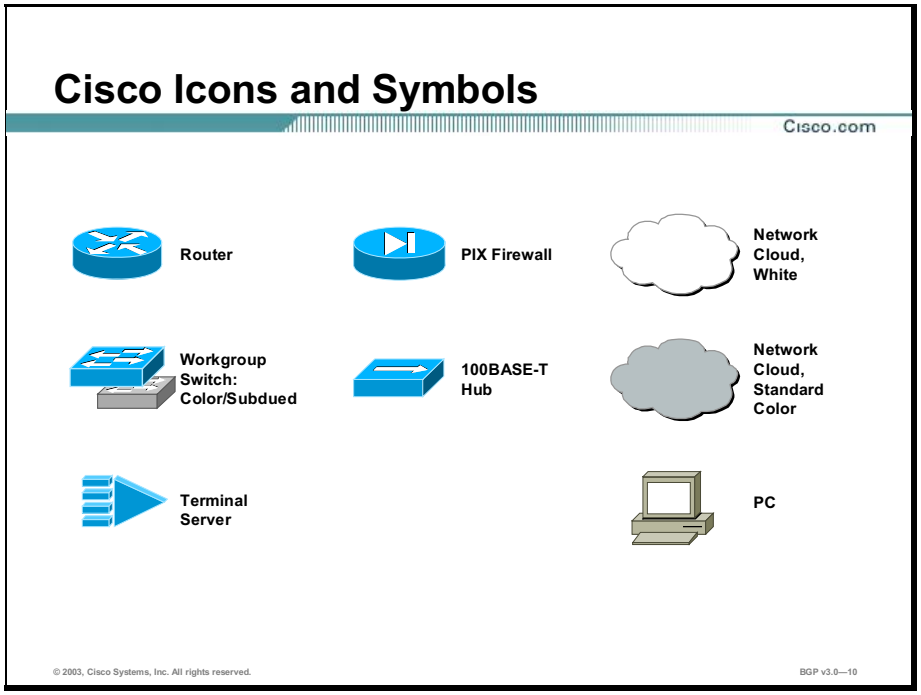


The schedule reflects the recommended structure for this course. This structure allows enough time for the instructor to present the course information and for you to work through the laboratory exercises. The exact timing of the subject materials and labs depends on the pace of your specific class.

 Copyright © 2003, Cisco Systems, Inc.

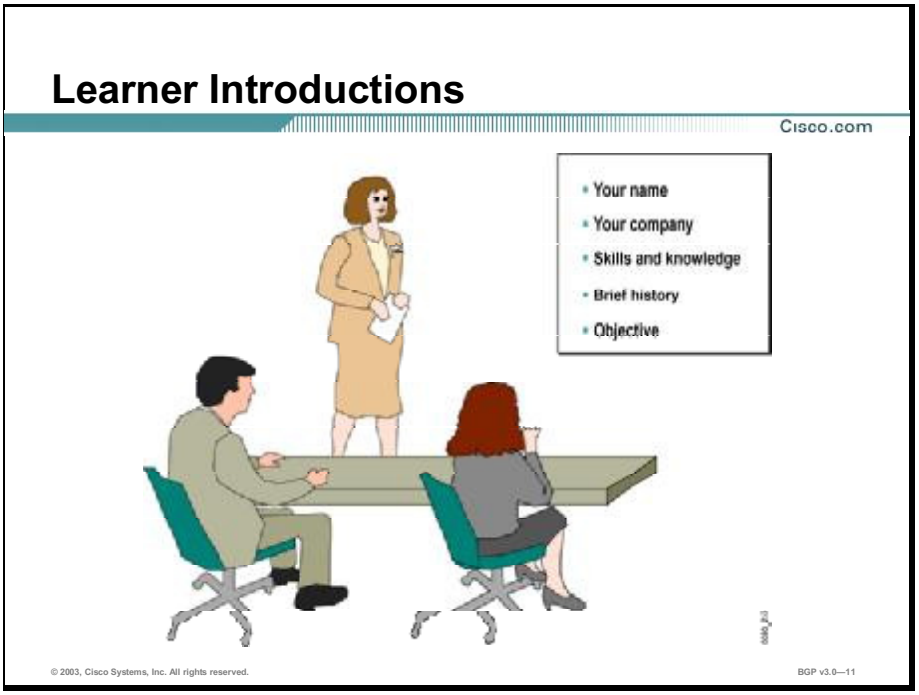
Icons and Symbols

This topic shows the Cisco icons and symbols used in this course.



Learner Introductions

This is the point in the course where you introduce yourself.



Prepare to share the following information:

- Your name
- Your company
- If you have most or all of the prerequisite skills
- A profile of your experience
- What you would like to learn from this course

 Copyright © 2003, Cisco Systems, Inc.

1

BGP Overview

Overview

The Border Gateway Protocol (BGP) is an interdomain (interautonomous system) routing protocol used to exchange routing information for the Internet. BGP, by design, is a very robust and scalable routing protocol. Because BGP is deployed as an interdomain routing protocol, it has many rich features that allow administrators to implement a variety of routing policies. This module covers basic BGP technology, details BGP session establishment and routing information exchange, and describes basic Cisco IOS® configuration and troubleshooting tasks.

Upon completing this module, you will be able to:

- Identify appropriate BGP usage and its limitations, given a network scenario
- Describe the concept of BGP neighbors and neighbor session establishment procedures, given a typical BGP network scenario
- Describe interdomain route processing, route propagation, and BGP path selection, given a diagram of an operational BGP network
- Successfully configure BGP, given a network consisting of multiple domains
- Verify proper operation and perform the steps necessary to correct basic BGP configuration errors, given a configured BGP network

Outline

The module contains these lessons:

- Introduction to BGP
- BGP Session Establishment
- BGP Route Processing
- Basic BGP Configuration
- Monitoring and Troubleshooting BGP

 Copyright © 2003, Cisco Systems, Inc.

Introduction to BGP

Overview

The Border Gateway Protocol (BGP) is an interautonomous system (AS) routing protocol used to exchange routing information between the Internet and Internet service providers (ISPs). BGP is a very robust and scalable routing protocol, as evidenced by the fact that it is the routing protocol employed on the Internet. This lesson introduces basic BGP characteristics and features.

Importance

Service providers and customer networks, such as universities and corporations, usually use an Interior Gateway Protocol (IGP) such as Routing Information Protocol (RIP), Enhanced Interior Gateway Routing Protocol (EIGRP), or Open Shortest Path First (OSPF) for the exchange of routing information within their networks. Any communication between these interior gateway protocols and the Internet or between service providers will be accomplished through BGP.

Objectives

Upon completing this lesson, you will be able to:

- Describe the requirements and use of scalable interdomain routing
- Identify basic BGP characteristics and features
- Explain why BGP is used by single-homed customers
- Explain why BGP is used by multihomed customers
- Explain why BGP is used in transit autonomous systems
- Explain BGP limitations

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Successful completion of *Building Scalable Cisco Internetworks* (BSCI) or equivalent

Outline

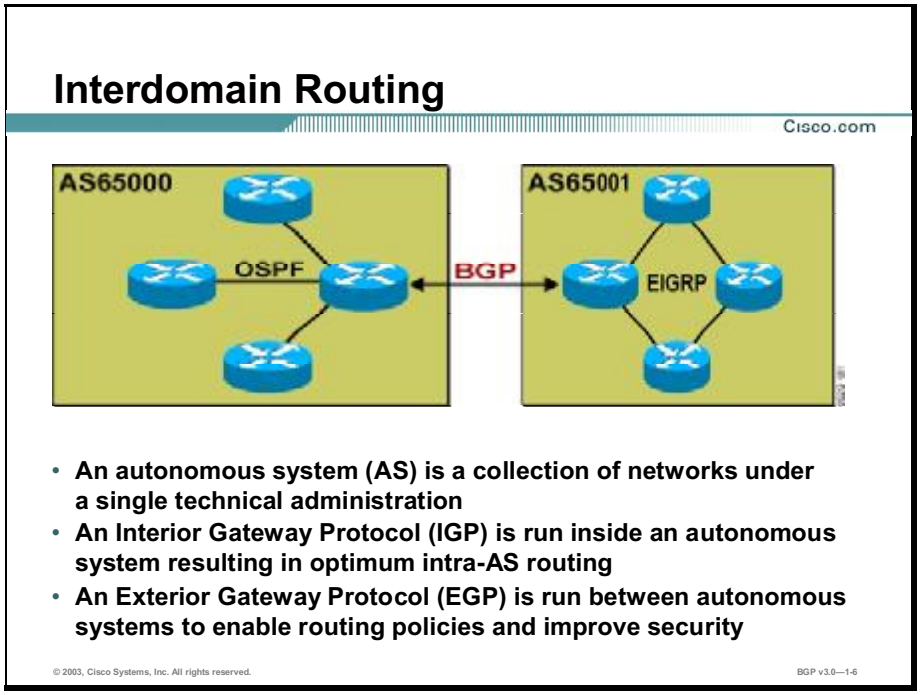
This lesson includes these topics:

- Overview
- Interdomain Routing
- BGP Characteristics
- Single-Homed Customers
- Multihomed Customers
- Transit Autonomous Systems
- BGP Limitations
- Summary
- Assessment (Quiz): Introduction to BGP

 Copyright © 2003, Cisco Systems, Inc.

Interdomain Routing

This topic defines interdomain routing and the design goals of interdomain routing protocols.

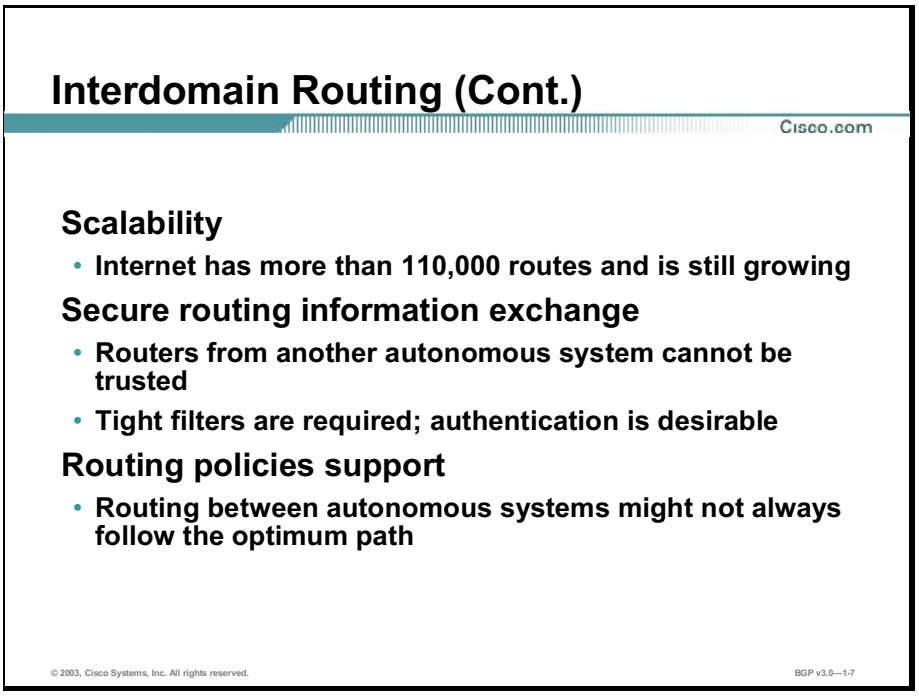

Interdomain Routing
Cisco.com
AS65000
OSPF
BGP
AS65001
EIGRP
• An autonomous system (AS) is a collection of networks under a single technical administration
• An Interior Gateway Protocol (IGP) is run inside an autonomous system resulting in optimum intra-AS routing
• An Exterior Gateway Protocol (EGP) is run between autonomous systems to enable routing policies and improve security
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-6

When talking to people involved with Internet routing, network administrators commonly use the terms autonomous system (AS), interdomain routing, and interior and exterior routing protocols. These terms, confusing for a novice, are defined below:

- An AS is a collection of networks under a single technical administration. Other definitions refer to a collection of routers or IP prefixes, but in the end they all mean the same entity. The important principle is the technical administration, which means sharing the same routing protocol and routing policy. Legal and administrative ownership of the routers does not matter in terms of autonomous systems. Autonomous systems are identified by AS numbers. AS numbers are 16-bit, unsigned integers ranging from 1 to 65535. Public AS numbers (1 – 64511) are assigned and managed by an Internet registry. A range of private AS numbers (64512 – 65535) has been reserved for customers that need an AS number to run BGP in their private networks.

- Interdomain routing is routing between autonomous systems. It is usually based on a set of policies, not just the technical characteristics of the underlying infrastructure.

- Exterior routing protocols (BGP being the only exterior routing protocol used today) are protocols that have the right set of functions to support various interdomain routing policies. Such protocols are contrary to interior routing protocols (for example, OSPF, RIP, or EIGRP), which only focus on finding the optimum (usually fastest) route between two points, with no respect to routing policies.

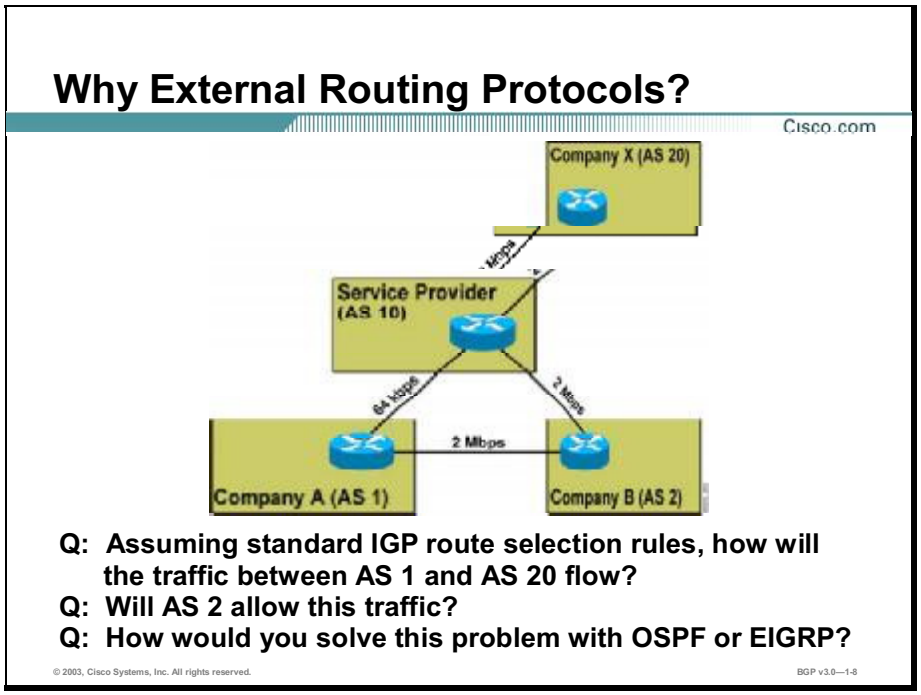

Interdomain Routing (Cont.)
Cisco.com
Scalability
• Internet has more than 110,000 routes and is still growing
Secure routing information exchange
• Routers from another autonomous system cannot be trusted
• Tight filters are required; authentication is desirable
Routing policies support
• Routing between autonomous systems might not always follow the optimum path
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-7

The design goals for any interdomain routing protocol include the following:

- Scalability: An interdomain routing protocol has to be able to support Internet routing, which consists of more than 110,000 routes.
- Secure information exchange: As the routers from other autonomous systems cannot be trusted, tight filters on routing updates and router authentication are desirable features.
- Support for routing policies: Routing between autonomous systems might not always follow the optimum path, and exterior routing protocols have to support a wide range of customer requirements.

 Copyright © 2003, Cisco Systems, Inc.

Example


Why External Routing Protocols?
Cisco.com
Company X (AS 20)
Service Provider
(AS 10)
64 kbps
2 Mbps
2 Mbps
Company A (AS 1)
Company B (AS 2)
Q: Assuming standard IGP route selection rules, how will the traffic between AS 1 and AS 20 flow?
Q: Will AS 2 allow this traffic?
Q: How would you solve this problem with OSPF or EIGRP?
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-8

The figure illustrates the need for an interdomain routing protocol. It depicts two companies that are connected to the Internet via leased lines of differing speed.

In routing protocols other than BGP, routing decisions are normally made to take advantage of the highest bandwidth available. Doing so would make traffic between AS 1 and AS 20 flow via AS 2. This situation is not desirable for AS 2, because it would allow the users in Company A to generate traffic on the Internet access line purchased and paid for by Company B.

Company B is unlikely to allow traffic from Company A to reach the Internet using the Company B access line. Company B, in fact, could create an access-list blocking all IP packets from AS 1 transmitted on the 2-Mbps serial line from Company B to the Internet. That action would create a black hole because Company A would send its packets to Company B and then Company B would drop them.

To avoid this situation, Company B must make sure that the packets from Company A, which are destined to the Internet, are never sent to Company B. Also, Company B must make sure that packets from the Internet destined for Company A, are never sent using the Internet access line to Company B. Company B could implement a routing policy, which indicates that AS 2 will receive reachability information from AS 1 for its own use, but AS 2 will not forward that particular information to the Internet. Also, AS 2 will receive reachability information about the Internet from its ISP but will never forward that information to AS 1. Only networks local to AS 2 are sent to AS 1.

The result of this routing policy would be that AS 1 sees all networks within AS 2 as reachable over the 2-Mbps link that directly connects AS 1 with AS 2. The routers in AS 1 will not see the rest of the Internet as reachable through AS 2. Therefore, AS 1 forwards packets toward the Internet directly over the 64-kbps link.

Also, the IP networks in AS 1 will appear reachable by AS 2 over the 2-Mbps link, which directly connects AS 1 with AS 2. However, the ISP will not receive that reachability information from AS 2; it will receive it only from AS 1. Therefore, traffic from the Internet to Company A will be transmitted over the 64-kbps link.

This routing policy is easy to implement when network administrators are using BGP, but impossible to implement with any other routing protocol. EIGRP, for example, can do route filtering only on individual IP subnets, not on all prefixes belonging to an AS. Link-state protocols, such as OSPF, cannot do powerful route filtering at all. BGP can do this routing based on AS numbers, which makes it possible to scale over the Internet.

Practice

Q1) What is an AS?

A) a set of Internet routers maintained by the IANA

B) a set of routing protocols designed to run inside a corporation or customer network

C) a set of protocols designed to implement routing policies

D) a collection of networks under a single technical administration

Q2) What are the three design goals of an interdomain routing protocol? (Choose three.)

A) secure routing information exchange

B) support for routing policies

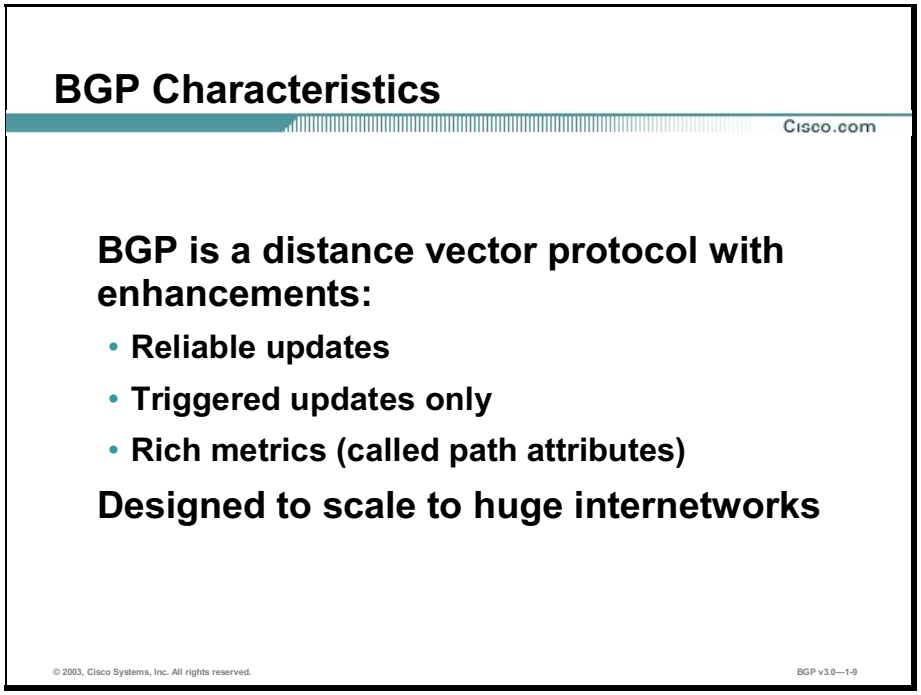
C) ability to provide WAN connectivity from remote customer sites

D) ability to scale to an extremely large number of routes

 Copyright © 2003, Cisco Systems, Inc.

BGP Characteristics

This topic describes the basic characteristics of BGP.

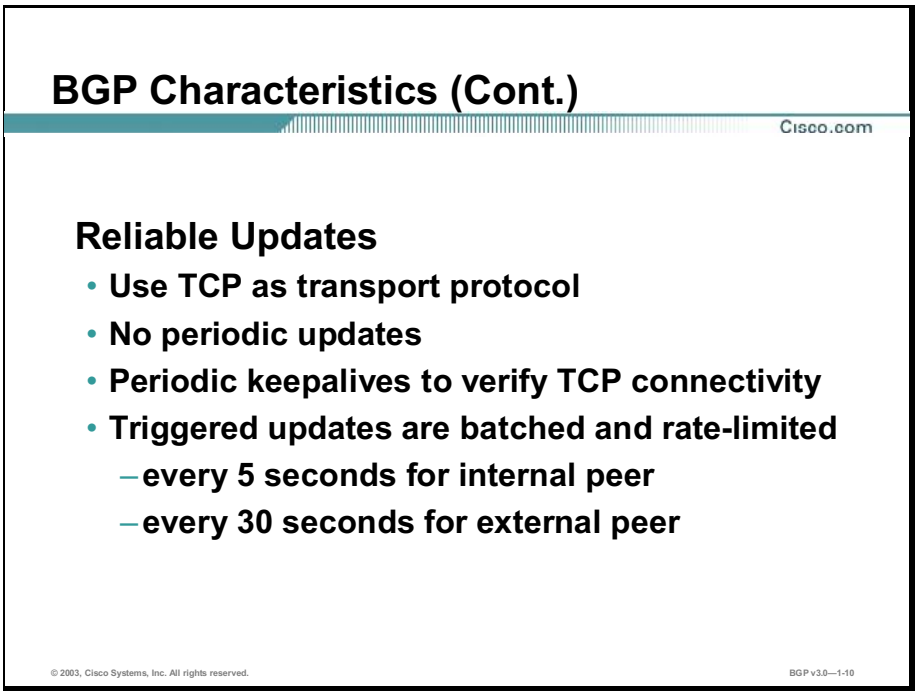

BGP Characteristics
Cisco.com
BGP is a distance vector protocol with enhancements:
• Reliable updates
• Triggered updates only
• Rich metrics (called path attributes)
Designed to scale to huge internetworks
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-9

BGP is a distance vector protocol. This means that it will announce to its neighbors those IP networks that it can reach itself. The receivers of that information will say “if that AS can reach those networks, then I can reach them via the AS.”

If two different paths are available to reach the same IP subnet, then the shortest path is used. This determination requires a mechanism capable of measuring the distance. All distance vector protocols have such mechanisms, called metrics. BGP contains a very sophisticated method of computing the shortest path by using attributes attached to the reachable IP subnet.

BGP sends routing updates to its neighbors by using a reliable transport. This means that the sender of the information always knows that the receiver has actually received it. As a result, there is no need for periodic updates or routing information refreshments. In BGP, only information that has changed is transmitted.

The reliable information exchange, combined with the batching of routing updates also performed by BGP, allows BGP to scale to Internet-sized networks.


BGP Characteristics (Cont.)
Cisco.com
Reliable Updates
• Use TCP as transport protocol
• No periodic updates
• Periodic keepalives to verify TCP connectivity
• Triggered updates are batched and rate-limited
– every 5 seconds for internal peer
– every 30 seconds for external peer
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-10

The reliable transport mechanism used by BGP is the standard TCP. BGP is an application protocol that uses the TCP and IP protocols for reliable connections.

Because BGP uses a reliable transport, the sender will know that the receiver has actually received the transmitted information. This capability makes periodic updates unnecessary.

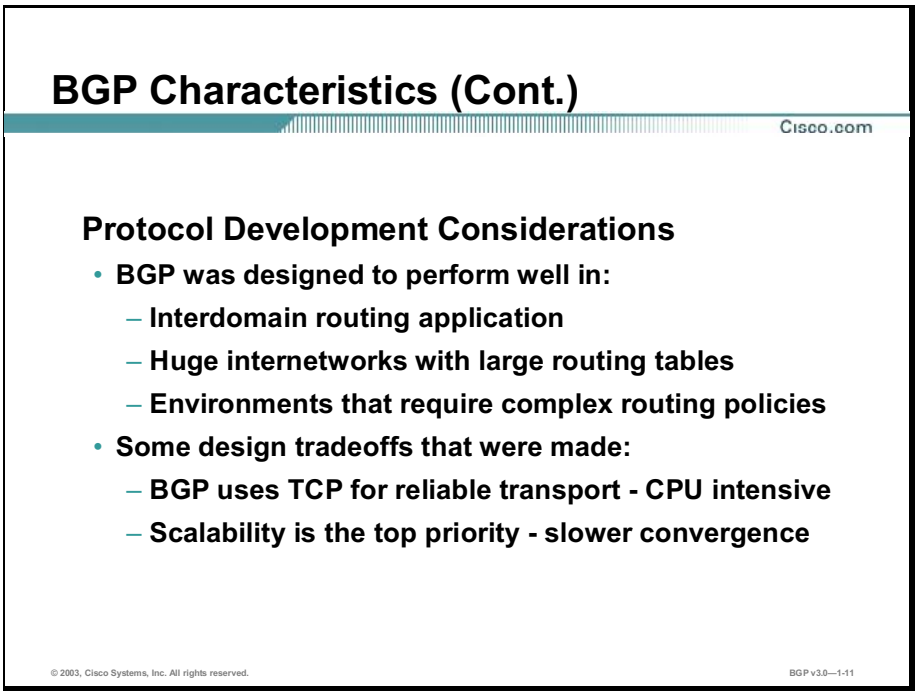
A router that has received reachability information from a BGP peer must be sure that the peer router is still there. Otherwise, the router could route traffic toward a next-hop router that is no longer available, causing the IP packets to be lost in a black hole. TCP does not provide the service to signal that the TCP peer is lost, unless some application data is actually transmitted between the peers. In an idle state, where there is no need for BGP to update its peer, the peer could be unreachable without TCP detecting it. Therefore, BGP takes care of detecting the presence of neighbors by periodically sending small BGP keepalive packets to them. These packets are considered application data by TCP and therefore must be transmitted reliably. According to the BGP specification, the peer router also must reply with a BGP keepalive packet.

When BGP was created, a key design goal was to be able to handle enormous amounts of routing information in very large and complex networks. In this environment, many links could go up and down (flapping), causing topology changes, which must be considered by the routing protocol. But low convergence time and quick responses to topology changes require fast updates and high CPU power to process both incoming and outgoing updates. The larger the network, the more updates per second can be expected if immediate response is required. The presence of too many updates in large networks can jeopardize network scalability.

 Copyright © 2003, Cisco Systems, Inc.

The designers of BGP decided that scalability was a more important issue than low convergence time, so BGP was designed to batch updates. Any changes received within the batch interval time are saved. At the end of the interval, only the remaining result is forwarded in an outgoing update. If a network flaps several times during the batch interval, only the state at the end of the interval is sent in an update. The batching feature avoids an uncontrolled flood of updates all over the Internet because the amount of updates is limited by the batching procedure.

Copyright © 2003, Cisco Systems, Inc.


BGP Characteristics (Cont.)
Cisco.com
Protocol Development Considerations
• BGP was designed to perform well in:
– Interdomain routing application
– Huge internetworks with large routing tables
– Environments that require complex routing policies
• Some design tradeoffs that were made:
– BGP uses TCP for reliable transport - CPU intensive
– Scalability is the top priority - slower convergence
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-11

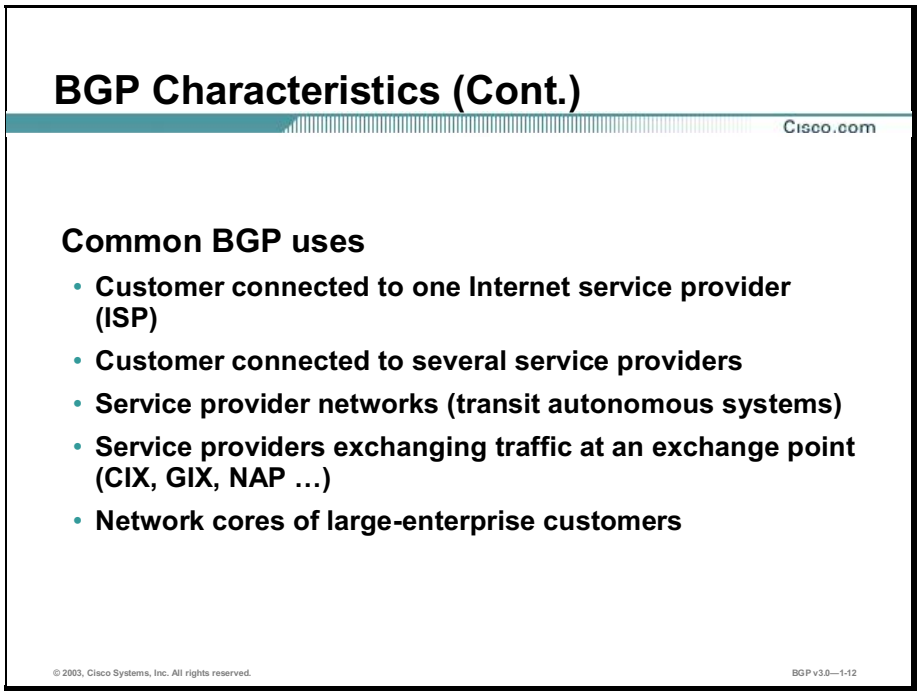
The designers of the BGP protocol have succeeded in creating a highly scalable routing protocol, which can forward reachability information between autonomous systems (also known as routing domains). BGP designers had to consider an environment with an enormous amount of reachable networks and complex routing policies driven by commercial rather than technical considerations.

TCP, a well-known and widely proven protocol, was chosen as the transport mechanism. That decision kept the BGP protocol simple, but it increased the CPU resource requirements for routers running BGP. The point-to-point nature of TCP might also introduce a slight increase in network traffic, because any update that should be sent to many receivers has to be multiplied into several copies, which are then transmitted on individual TCP sessions to the receivers.

Whenever there was a design choice between fast convergence and scalability, scalability was the top priority. Batching of updates and the relatively low frequency of keepalive packets are examples of where designers placed convergence time second to scalability.

Note BGP convergence times can be modified with the configuration of nondefault values for BGP scan and advertisement timers. Refer to the Optimizing BGP Scalability module for more information on tuning BGP convergence.

 Copyright © 2003, Cisco Systems, Inc.


BGP Characteristics (Cont.)
Cisco.com
Common BGP uses
• Customer connected to one Internet service provider (ISP)
• Customer connected to several service providers
• Service provider networks (transit autonomous systems)
• Service providers exchanging traffic at an exchange point (CIX, GIX, NAP …)
• Network cores of large-enterprise customers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-12

The figure shows typical scenarios where BGP is usable. These scenarios include the following:

- Customers connected to one or more service providers
- ISP networks themselves acting as transit systems and forwarding external traffic
- Very large enterprises using BGP as their core routing protocol

Practice

Q1) Which transport mechanism is used to exchange BGP routing updates?

A) UDP

B) TCP

C) SDP

D) LDAP

Q2) To achieve reliability and scalability, what BGP design tradeoff was made?

A) Route authentication is not supported.

B) Convergence is slower due to batched updates and long keepalive timers.

C) BGP supports only specific IGPs.

D) BGP has to perform periodic updates to synchronize route tables.

Q3) What are BGP keepalive messages used for?

A) to form the initial adjacency between BGP peers

B) to calculate the best path by determining the round-trip delay

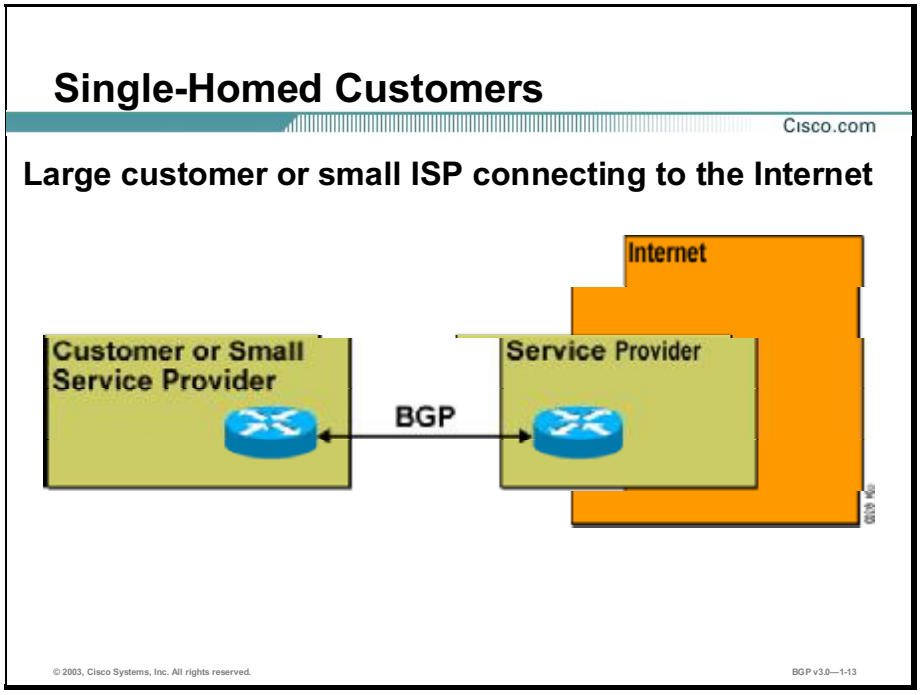
C) to detect a neighbor when the BGP session is in the idle state

D) to acknowledge the receipt of BGP route updates from the BGP neighbor

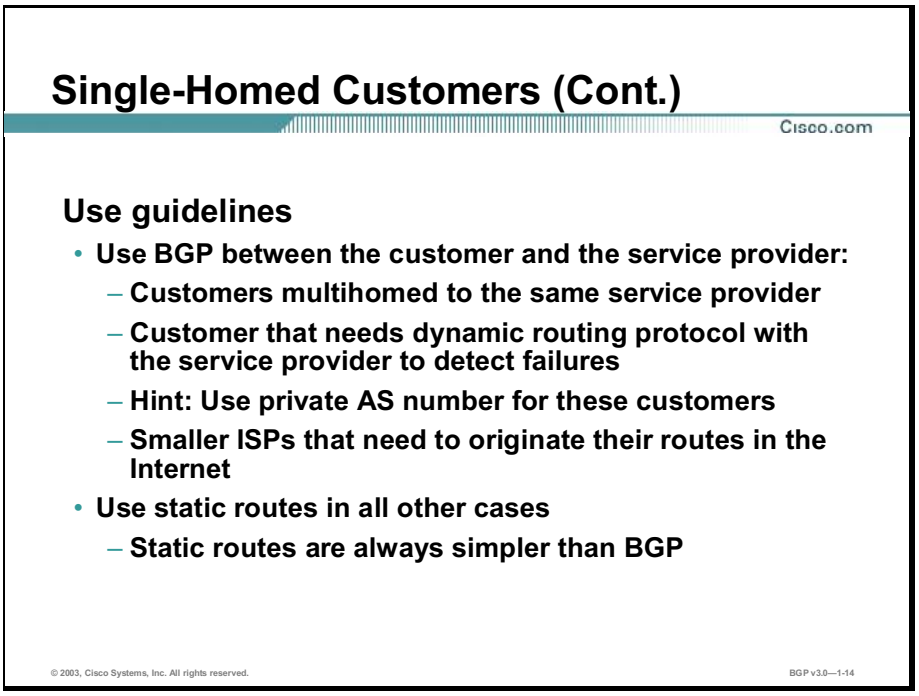
 Copyright © 2003, Cisco Systems, Inc.

Single-Homed Customers

This topic explains when a single-homed customer should use BGP as an interdomain routing protocol.



The figure shows a customer network connected to the Internet using a single ISP, but such a scenario is generally not the case when BGP is used. Normal Internet access to a single ISP does not require BGP; static routes are more commonly used to handle this situation. Small ISPs buying Internet connectivity from other ISPs use this type of connectivity more often, especially if they want to start their business the proper way—by using their own AS number and having their own address space.


Single-Homed Customers (Cont.)
Cisco.com
Use guidelines
• Use BGP between the customer and the service provider:
– Customers multihomed to the same service provider
– Customer that needs dynamic routing protocol with the service provider to detect failures
– Hint: Use private AS number for these customers
– Smaller ISPs that need to originate their routes in the Internet
• Use static routes in all other cases
– Static routes are always simpler than BGP
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-14

Under certain conditions, BGP must be configured between the customer and the service provider. For example, BGP is needed when customers are multihomed to the same service provider (that is, the customer networks have multiple links connecting them with the service provider network) and require dynamic routing protocol interaction with the service provider to detect link failures. Private AS numbers (AS numbers above 64512) are usually implemented in BGP configurations for these customers.

Customers that plan to connect to more than one ISP, and small ISPs that plan to have multiple Internet connections in the future, usually use BGP with their service provider. They use this option even when they have a single link with the service provider in order to be prepared for future upgrades.

In all other cases, using static routes from the service provider toward the customer and using a default static route from the customer toward the service provider is the preferred method of provider-to-customer routing in the Internet.

Practice

Q1) What are the two reasons to use BGP between a single-homed customer and a service provider? (Choose two.)

A) if multiple connections to the same provider will be used

B) if the customer is using RFC 1918 IP addressing

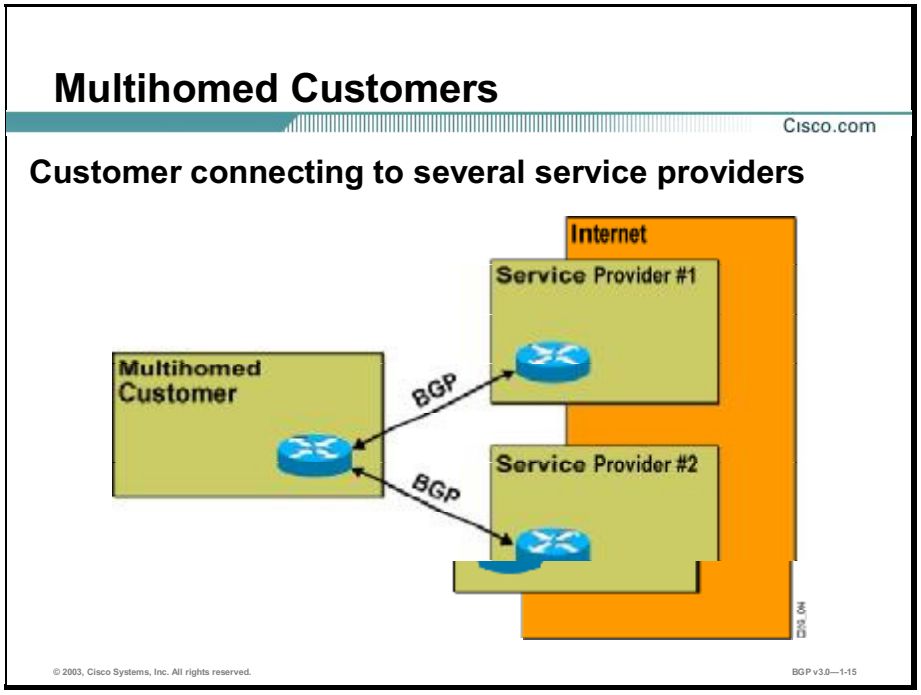
C) as a preparation step for a future upgrade to multiple providers

D) if the customer has a large WAN with a single Internet connection

 Copyright © 2003, Cisco Systems, Inc.

Multihomed Customers

This topic explains when BGP is appropriate for the multihomed customer. BGP use guidelines for multihoming are also discussed.



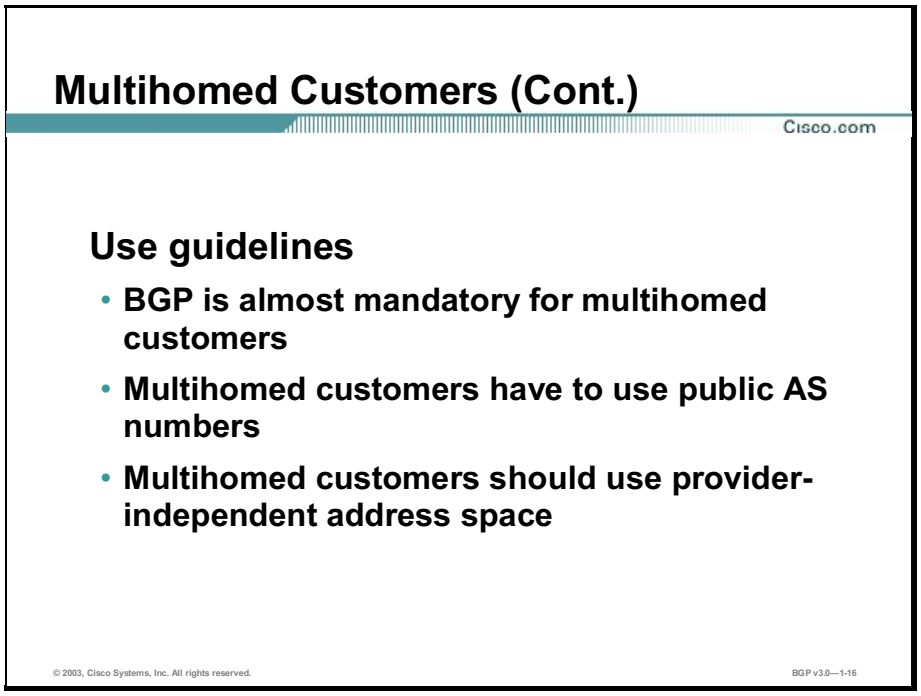
The figure illustrates a customer network connected to two different ISPs, which requires the use of BGP for full redundancy.

The customer must have its own officially assigned AS number. The customer is also responsible for announcing its own IP networks to both ISPs. Both ISPs forward all routes received from the Internet to the customer network. The customer should avoid forwarding any routing information received from one ISP to the other. Otherwise, the customer will become a transit provider between the two ISPs. This is a situation that most customers like to avoid because it creates a resource drain on routers and network links.

Full redundancy is achieved in this setup. If either of the two access links fails, the reachability information previously transmitted on the now-failed link will be withdrawn. But BGP reachability information is still announced by the customer's router over the remaining link. So the ISP will still see all networks within the customer AS as reachable but only over the remaining path. Also, received routes from the Internet will be withdrawn when the link fails, but routes received over the remaining link are not affected. So the Internet, including the ISP to which the direct connection has failed, is still reachable over the remaining link.

This design can also handle other problems. A case where both access links are available, but the connection between one of the ISPs and the rest of the Internet is lost, works as follows: The ISP that has a problem reaching the rest of the Internet withdraws all those routes and tells the customer AS that it can no longer reach the Internet. But the networks local to the ISP with the Internet reachability problem are still reachable by the customer, so those routes are not withdrawn. The networks in the customer AS are still reachable by the ISP in trouble, but that ISP can no longer forward the announcement to the rest of the Internet. The rest of the Internet will, however, see the customer networks reachable over the path to the other ISP, which is fully functional.

 Copyright © 2003, Cisco Systems, Inc.


Multihomed Customers (Cont.)
Cisco.com
Use guidelines
• BGP is almost mandatory for multihomed customers
• Multihomed customers have to use public AS numbers
• Multihomed customers should use provider-independent address space
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-16

The following use guidelines apply to multihomed customers:

- Although there are designs where BGP could be avoided, most multihomed customers need to use BGP with their service providers.
- The multihomed customers must have their own AS number, and it is recommended to use a public AS number.
- Multihomed customers should use a provider-independent address space, which is allocated to them directly by an Internet registry.

Practice

Q1) Why should the multihomed customer avoid forwarding any routing information received from one ISP to the other?

A) to guarantee that routing loops will not occur

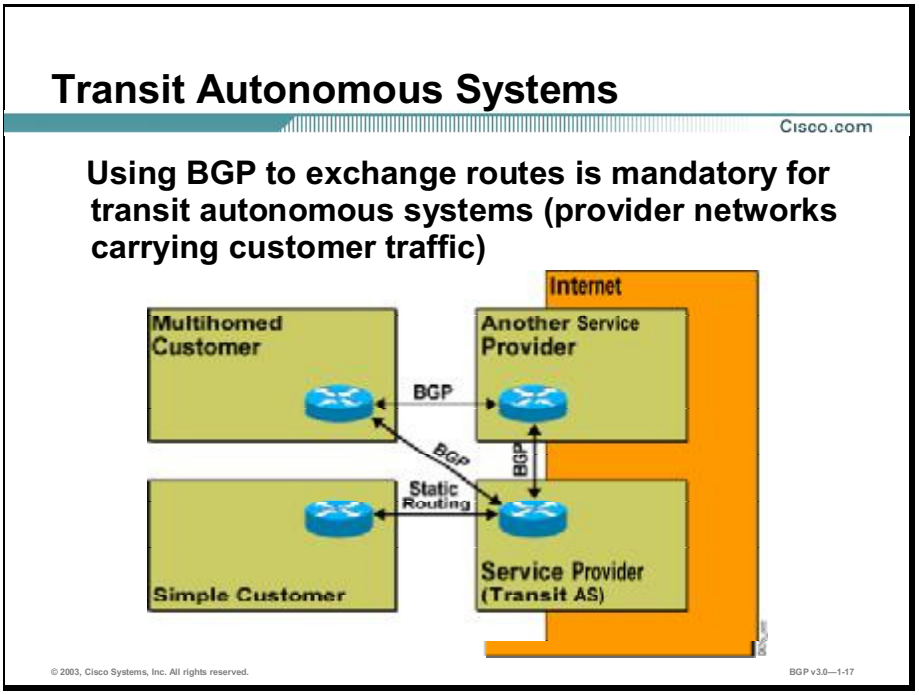
B) to avoid the possibility that the customer could become a transit backbone between the two providers

C) to prevent route leaking of the customer IGP routes into the Internet

D) to prevent denial-of-service attacks from the Internet

Transit Autonomous Systems

This topic describes the use of BGP in a transit autonomous backbone.



BGP is most commonly implemented in service provider networks to ensure connectivity between their customers and the rest of the Internet. An ISP might exchange BGP updates with the customers or use static routing toward them. It also connects to other ISPs and is required to forward the routes received from the customers to the rest of the Internet, as well as in the other direction. As a result, user data traffic starts to flow between the customers and the rest of the Internet. Such a network, providing transit services to traffic originated in other networks, is thus called a transit autonomous system. A transit AS is an AS that exchanges BGP routing information with other autonomous systems and forwards information received from one AS to another AS.

When routing information is forwarded, the receiver will see an available path to a destination and start transmitting user data toward the destination using that path. The transit AS must be prepared to relay the user data, as explained later in this course.

ISP networks could have dedicated peer-to-peer connections, using, for example, packet over SONET (POS). These connections are sometimes called private peering. ISPs also interconnect at exchange points. Technically, an exchange point is just a multiaccess subnet: a LAN (for example, a Gigabit Ethernet or Fast Ethernet switch), a Dynamic Packet Transport (DPT) ring, or an ATM switch. Many ISPs can connect to an exchange point and establish BGP sessions.

The benefit of an exchange point is that it is highly scalable. There is no need for additional physical interfaces in the ISP border router, when a new ISP is launched. If the already-established ISPs want to, they can open a BGP session with the new ISP. When this is done, they start to exchange routing information and then user data traffic over the exchange point.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What is the main function of a transit AS in BGP?

A) to facilitate customer connectivity to the Internet

B) to forward routing information learned from one AS into another AS

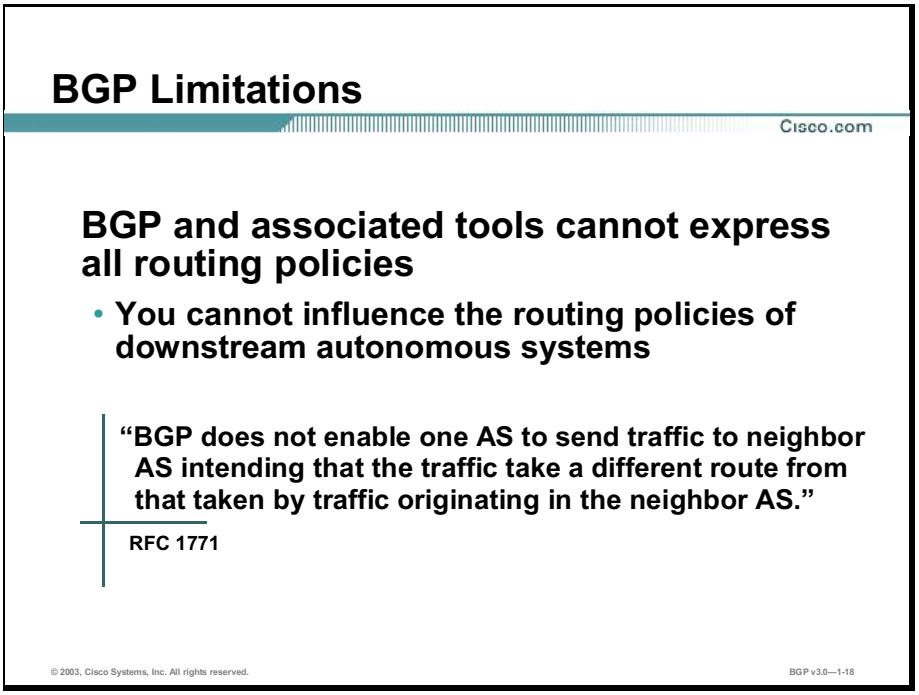
C) to direct routing policy in neighboring autonomous systems

D) to allow non-BGP configured customer routers to connect to the Internet

Copyright © 2003, Cisco Systems, Inc.

BGP Limitations

This topic describes some of the limitations of BGP.


BGP Limitations
Cisco.com
BGP and associated tools cannot express all routing policies
• You cannot influence the routing policies of downstream autonomous systems
"BGP does not enable one AS to send traffic to neighbor AS intending that the traffic take a different route from that taken by traffic originating in the neighbor AS."
RFC 1771
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-18

BGP enabled routers make forwarding decisions based on the destination IP address only; the source IP address does not affect the decision. If an AS acts as a transit AS for other autonomous systems, the IP packets created and transmitted from the other autonomous systems are not treated differently from the IP packets created and transmitted from the local AS. If the local AS has decided that the best path to reach a certain destination is via a specific next-hop router, then it will route all user data traffic toward the final destination via that specific next-hop router. The local AS makes its decision based on destination address only, regardless of which IP host has sourced the IP packets.

Practice

Q1) What are BGP forwarding decisions based upon?

A) source IP address

B) destination IP address

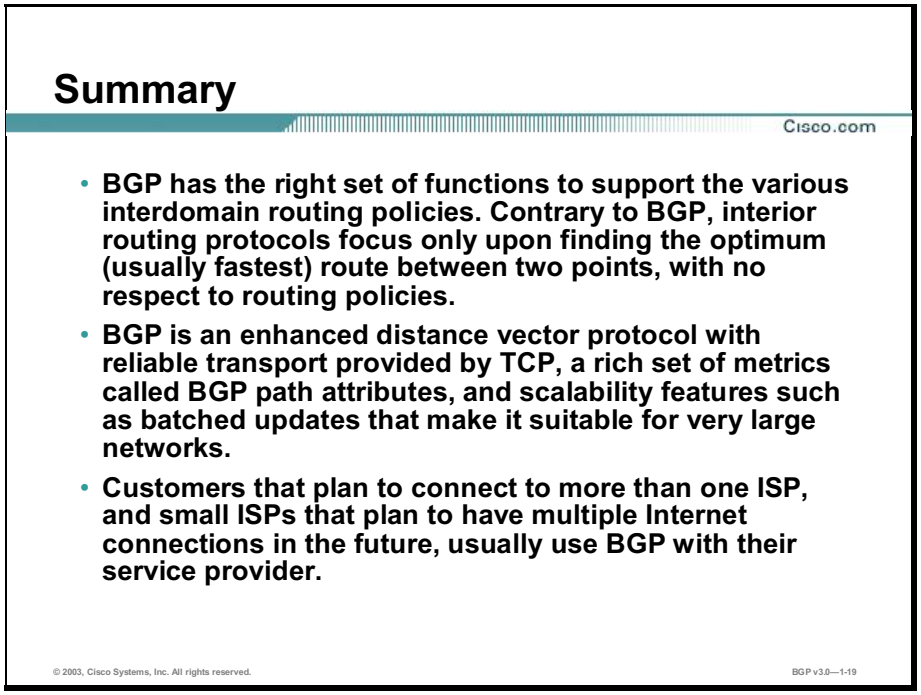
C) next-hop attribute

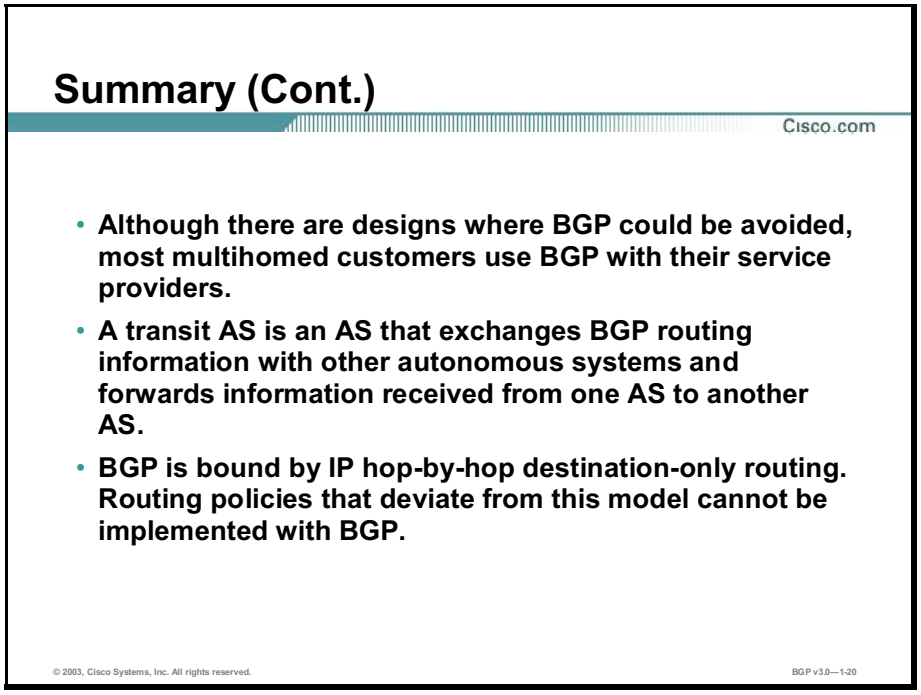
D) remote AS policies

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• BGP has the right set of functions to support the various interdomain routing policies. Contrary to BGP, interior routing protocols focus only upon finding the optimum (usually fastest) route between two points, with no respect to routing policies.
• BGP is an enhanced distance vector protocol with reliable transport provided by TCP, a rich set of metrics called BGP path attributes, and scalability features such as batched updates that make it suitable for very large networks.
• Customers that plan to connect to more than one ISP, and small ISPs that plan to have multiple Internet connections in the future, usually use BGP with their service provider.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-19


Summary (Cont.)
Cisco.com
• Although there are designs where BGP could be avoided, most multihomed customers use BGP with their service providers.
• A transit AS is an AS that exchanges BGP routing information with other autonomous systems and forwards information received from one AS to another AS.
• BGP is bound by IP hop-by-hop destination-only routing. Routing policies that deviate from this model cannot be implemented with BGP.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-20

Next Steps

After completing this lesson, go to:

- BGP Session Establishment lesson

References

For additional information, refer to these resources:

- For more information on basic BGP, refer to “Border Gateway Protocol” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

- For further information on basic BGP, refer to “Configuring BGP” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121cgcr/ip_c/ipcprt2/1cdbgp.htm

 Copyright © 2003, Cisco Systems, Inc.

Quiz: Introduction to BGP

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe the requirements and use of scalable interdomain routing
- Identify basic BGP characteristics and features
- Explain why BGP is used by single-homed customers
- Explain why BGP is used by multihomed customers
- Explain why BGP is used in transit autonomous systems
- Explain BGP limitations

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What three items are BGP enhancements to traditional distance vector routing protocols? (Choose three.)

A) reliable updates

B) use of triggered updates only

C) enhanced security

D) rich metrics

Copyright © 2003, Cisco Systems, Inc.

Q2) What protocol facilitates reliable update capabilities in BGP?

A) TCP

B) UDP

C) HSRP

D) ICMP

Q3) What are three characteristics of an AS? (Choose three.)

A) uses IGPs for intradomain routing

B) uses EGPs for interdomain routing

C) is a collection of networks under a common administrative authority

D) consists of a group of network domains

Q4) What three scenarios are common where BGP is used? (Choose three.)

A) a customer with a connection to multiple service providers

B) service provider networks acting as transit systems and forwarding external traffic through their network

C) single-site customer intranet with complex administrative policies between departments

D) as the core routing protocol in very large enterprise networks

Q5) What are three recommended BGP use guidelines for multihomed customer networks? (Choose three.)

A) Most multihomed customers should use BGP with their service providers.

B) Most multihomed customers should forward routing information received from one provider to the other provider.

C) The multihomed customer must have its own public AS number.

D) Multihomed customers should use a provider-independent, public address space.

 Copyright © 2003, Cisco Systems, Inc.

Q6) What is a limitation of the BGP routing protocol?

A) You cannot use BGP to implement hop-by-hop routing policy controls.

B) You cannot use BGP to influence the routing policy in a downstream AS.

C) BGP cannot control forwarding of packets based on their destination address.

D) BGP cannot scale to very large networks with more than 110,000 routes.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

BGP Session Establishment

Overview

Border Gateway Protocol (BGP) is an exterior gateway protocol (EGP) designed for scalability and policy control. As a result, BGP requires neighboring routers to be explicitly configured before BGP routing updates can be sent between them. This situation differs from Interior Gateway Protocols (IGPs) such as Enhanced Interior Gateway Routing Protocol (EIGRP) and Open Shortest Path First (OSPF) that discover neighbors through the use of a broadcast packet or a hello protocol. In this lesson, BGP neighbor session establishment procedures are discussed.

Importance

Understanding the BGP neighbor session establishment process is a key component to understanding the fundamental operation of the BGP protocol. It also forms a knowledge base for later lessons, including configuring basic BGP.

Objectives

Upon completing this lesson, you will be able to:

- Describe how BGP discovers neighbors
- Explain BGP session establishment procedures
- Explain the role of the BGP keepalive in session establishment and maintenance
- Describe how optional MD5 authentication can protect sessions between BGP peers

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Successful completion of *Building Scalable Cisco Internetworks* (BSCI) or equivalent

Outline

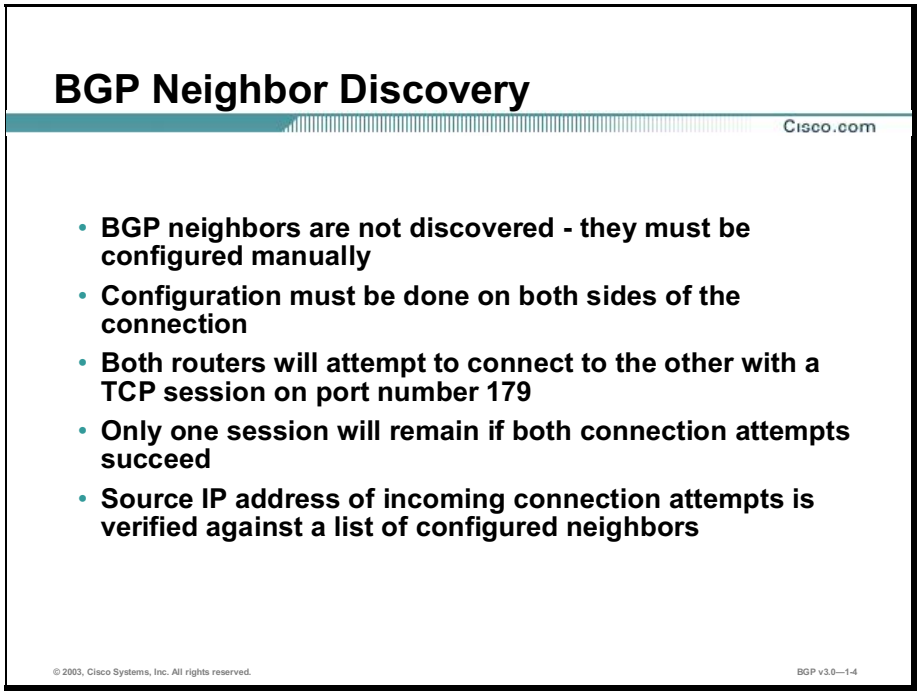
This lesson includes these topics:

- Overview
- BGP Neighbor Discovery
- Establishing a BGP Session
- BGP Keepalives
- MD5 Authentication
- Summary
- Assessment (Quiz): BGP Session Establishment

 Copyright © 2003, Cisco Systems, Inc.

BGP Neighbor Discovery

This topic describes how the BGP routing protocol discovers neighbors.


BGP Neighbor Discovery
Cisco.com
• BGP neighbors are not discovered - they must be configured manually
• Configuration must be done on both sides of the connection
• Both routers will attempt to connect to the other with a TCP session on port number 179
• Only one session will remain if both connection attempts succeed
• Source IP address of incoming connection attempts is verified against a list of configured neighbors
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-4

Unlike other routing protocols, BGP has no means of automatically detecting neighbors. The BGP protocol is carried in a TCP session, which must be opened from one router to the other. In order to do so, the router attempting to open the session must be manually configured with neighbor information indicating to which IP address to direct its connection attempts.

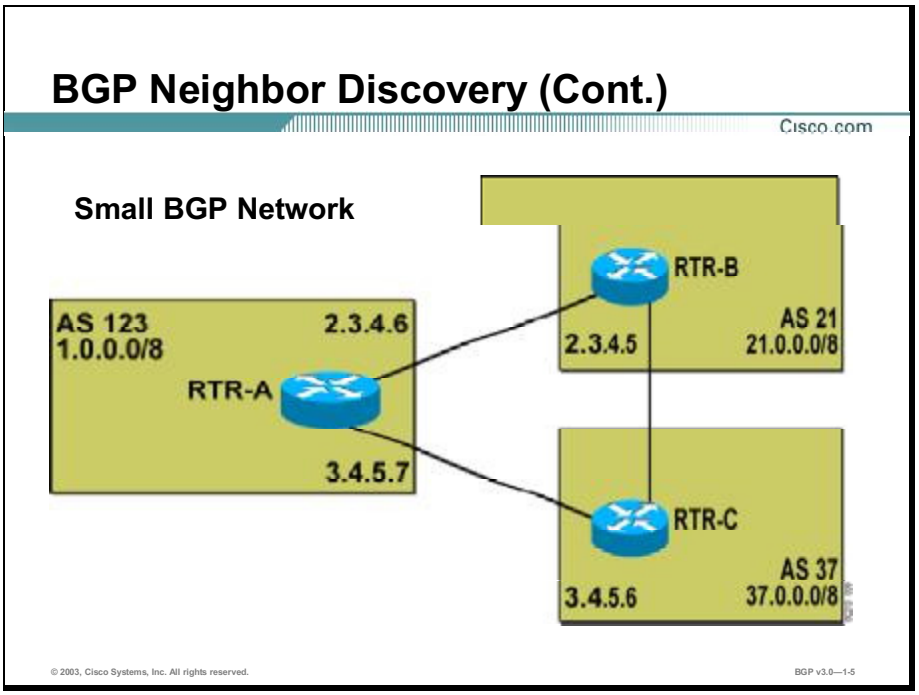
The router that receives the incoming connection attempts does not answer them if the attempts are not from one of the configured neighbors. The IP source address of the connection attempt packet (TCP SYN packet) is verified against the list of IP addresses that the router itself would direct its connection attempts to.

In order to succeed in the connection attempts, both routers are required to be configured to reach each other. A side effect of this situation is that they will both attempt to connect. This side effect adds robustness to the session establishment process, but it also introduces the risk that two BGP sessions will be established between a pair of BGP routers.

Two routers should have only a single BGP session between them. The router-ID values exchanged when the BGP session is established allow the BGP router to detect when two parallel sessions exist. Only the session initiated by the router with the numerically higher router-ID will be retained. The other session is dropped.

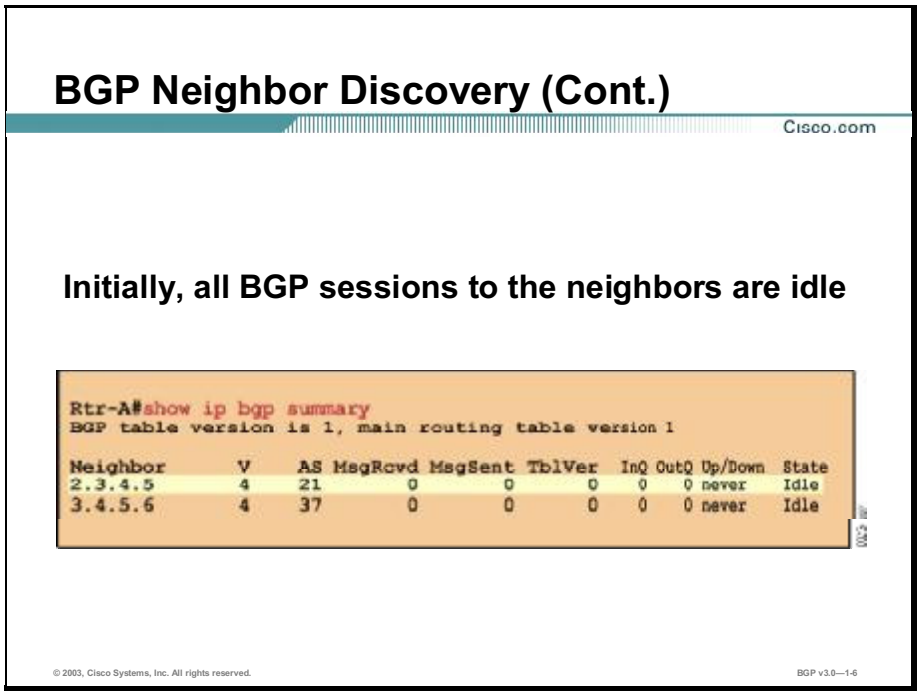
A router may not open a BGP session to itself. If the configured neighbor IP address is, in fact, an IP address of the local router, the router will recognize the problem and tear down the session. The router-ID is also used for this verification.

Example



The network displayed in the figure serves as the sample network to generate printouts in the following examples.

 Copyright © 2003, Cisco Systems, Inc.


BGP Neighbor Discovery (Cont.)
Cisco.com
Initially, all BGP sessions to the neighbors are idle
Rtr-A#show ip bgp summary
BGP table version is 1, main routing table version 1
Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State
2.3.4.5 4 21 0 0 0 0 0 never Idle
3.4.5.6 4 37 0 0 0 0 0 never Idle
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-6

The **show ip bgp summary** command gives an overview of the BGP status. Each configured neighbor is listed in the output of the command. The IP address to which the connection attempts are directed is also displayed, along with BGP version number, the remote AS number, some counter values, the status of the session, and how long ago the session changed state.

The “Idle” state indicates that the router is currently not attempting any connection establishments.

The different states for a BGP connection are Idle, Active, OpenSent, OpenConfirm, and Established.

Practice

Q1) How are BGP neighbors discovered?

A) Through the use of broadcast packets.

B) Through the use of IP multicast packets.

C) Through the use of a hello protocol.

D) Neighbors are not discovered but instead must be manually configured.

Q2) What protocol and port are used by BGP for session establishment? (Choose two.)

A) TCP

B) UCP

C) 53

D) 161

E) 179

Q3) What is the state of a BGP router prior to forming a BGP neighbor session?

A) Down

B) Idle

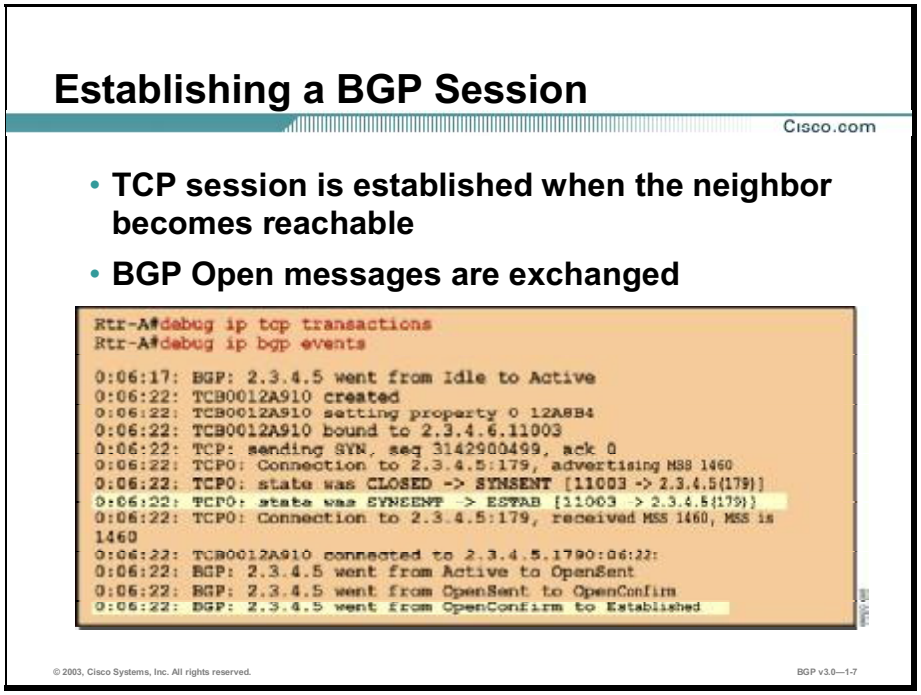
C) Not connected

D) Unestablished

 Copyright © 2003, Cisco Systems, Inc.

Establishing a BGP Session

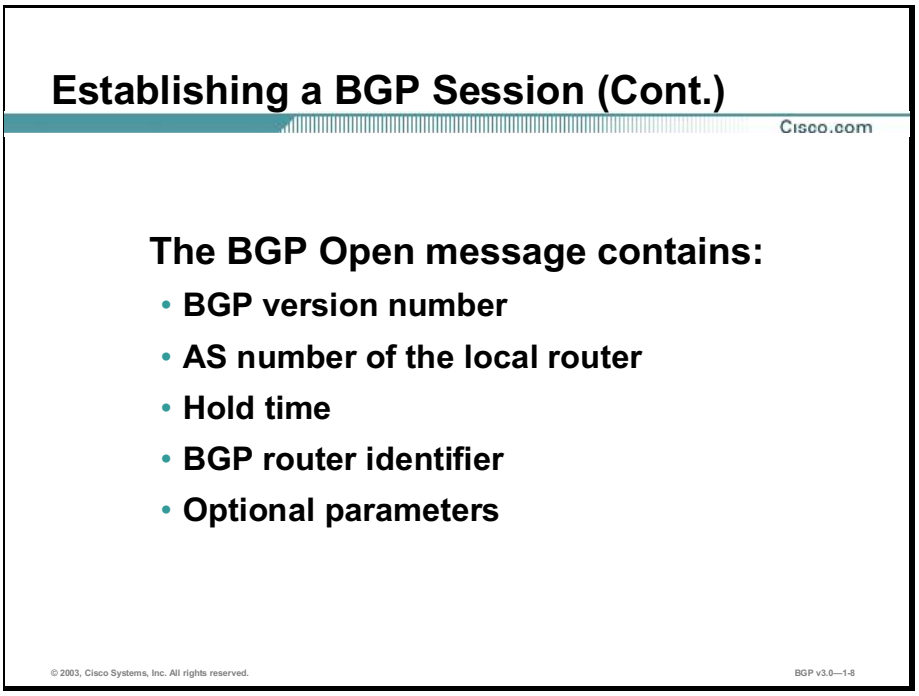
This topic explains the BGP session establishment process.


Establishing a BGP Session
Cisco.com
• TCP session is established when the neighbor becomes reachable
• BGP Open messages are exchanged
Rtr-A#debug ip tcp transactions
Rtr-A#debug ip bgp events
0:06:17: BGP: 2.3.4.5 went from Idle to Active
0:06:22: TCB0012A910 created
0:06:22: TCB0012A910 setting property 0 12A8B4
0:06:22: TCB0012A910 bound to 2.3.4.6.11003
0:06:22: TCP: sending SYN, seq 3142900499, ack 0
0:06:22: TCP0: Connection to 2.3.4.5:179, advertising MSS 1460
0:06:22: TCP0: state was CLOSED -> SYNSENT [11003 -> 2.3.4.5(179)]
0:06:22: TCP0: state was SYNSENT -> ESTAB [11003 -> 2.3.4.5(179)]
0:06:22: TCP0: Connection to 2.3.4.5:179, received MSS 1460, MSS is 1460
0:06:22: TCB0012A910 connected to 2.3.4.5.179 0:06:22:
0:06:22: BGP: 2.3.4.5 went from Active to OpenSent
0:06:22: BGP: 2.3.4.5 went from OpenSent to OpenConfirm
0:06:22: BGP: 2.3.4.5 went from OpenConfirm to Established
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-7

Before any connection attempt is made, the BGP peer relation must have left the Idle state and entered the Active state. For a BGP session between two routers in different autonomous systems, this situation happens when the IP address of the remote router becomes reachable on a directly connected interface.

The debug output shows how the router creates a socket data structure and binds it to its local IP address 2.3.4.6 and a high port number 11003. Then the router sends a TCP SYN packet to the configured peer router IP address of 2.3.4.5 and the well-known destination port 179. The connection attempt succeeds, and the TCP session is now ready to transfer the BGP information.

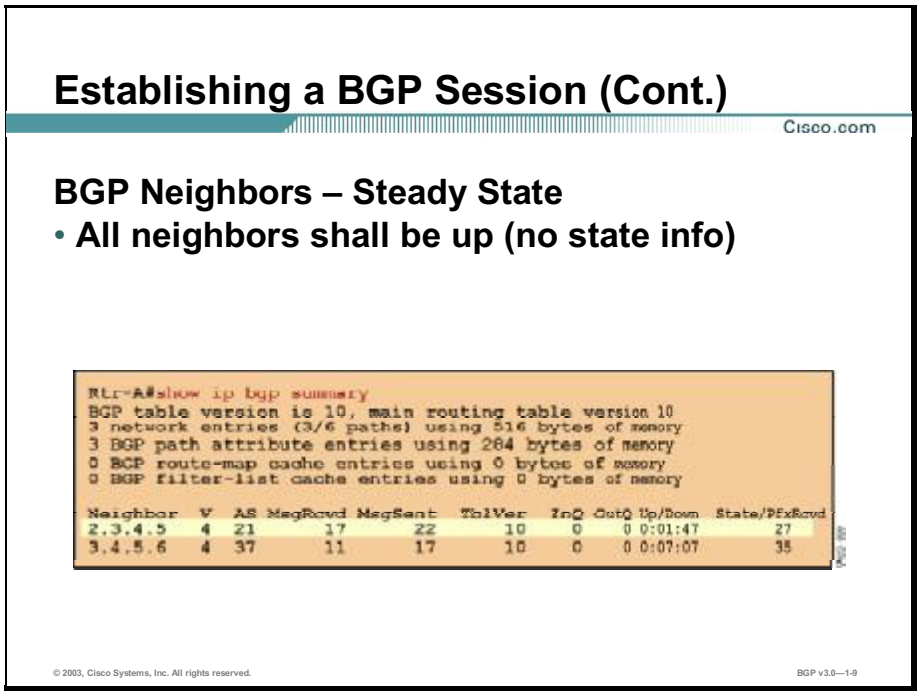
The first BGP information sent is the BGP Open message. The BGP session now goes from Active state to OpenSent state while waiting for the other router to respond. If the peer router accepts the parameters in the Open message, it responds with its own Open message. When the local router receives this message, the state goes from OpenSent to OpenConfirm. The local router now verifies the peer router parameters in its Open message. If they are accepted, a keepalive packet is sent to signal this. The state is now “Established.”


Establishing a BGP Session (Cont.)
Cisco.com
The BGP Open message contains:
• BGP version number
• AS number of the local router
• Hold time
• BGP router identifier
• Optional parameters
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-8

The parameters in the BGP Open message are:

- **Version number:** the suggested version number. The highest common version that both routers support will be used. Most BGP implementations today use BGP version 4.
- **AS number:** the AS number of the local router. The peer router will verify this information. If it is not the AS number that is expected, the BGP session is torn down.
- **Hold time:** the number of seconds that may elapse between receptions of successive BGP messages. If the time is exceeded, the peer will be considered dead. The two routers will agree to use the lowest suggested value. When the session is established, both routers will use keepalive messages to make sure that the hold timer does not expire. A suggested hold-timer value of 0 indicates that the timer never expires and no keepalives should be sent.
- **BGP identifier:** a number uniquely identifying the router. The Cisco router will use one of its IP addresses for this, the router-ID. This router-ID is selected as the numerically highest IP address of any loopback interface. If there is no loopback interface, it will use the highest IP address of any interface being up at the time of starting the BGP process.
- **Optional parameters:** are type, length, value (TLV) encoded. An example of optional parameters is session authentication.

 Copyright © 2003, Cisco Systems, Inc.


Establishing a BGP Session (Cont.)
Cisco.com
BGP Neighbors – Steady State
• All neighbors shall be up (no state info)
Rtr-A#show ip bgp summary
BGP table version is 10, main routing table version 10
3 network entries (3/6 paths) using 516 bytes of memory
3 BGP path attribute entries using 264 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcvd
2.3.4.5 4 21 17 22 10 0 0 0:01:47 27
3.4.5.6 4 37 11 17 10 0 0 0:07:07 35
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-9

After the BGP sessions are in the Established state, routing information exchange can take place. The **show ip bgp summary** command output here indicates that a session is established by not displaying any information at all in the state column.

The counter values show how many messages have been received and sent on the session. InQ shows how many messages have been received but not yet processed. A high InQ number indicates lack of CPU resources to process the input. OutQ shows how many outgoing messages are queued. A high OutQ number indicates lack of bandwidth to transmit the outgoing messages or CPU overload of the other router.

TblVer (table version) is used by the BGP router to track the changes that need to be sent to the neighbors. There is a major table version number for the local BGP table. The table version number is displayed on the first line of output from this **show** command. There is also one table version number maintained for each of the neighbors, which is displayed on the information line of the neighbors.

Whenever a BGP router enters a change into its BGP table, the major table version number is incremented and the changed route is tagged with this number. When the time comes to update a specific neighbor, the router scans the BGP table and all changes it finds where the version number is between the neighbor version and the current table version are sent to the BGP neighbor in a single BGP routing update. After the entire table is scanned and all changes have been sent to the neighbor, the table version number of the neighbor is set to the highest value of the routes being sent.

A table version of a neighbor that is lower than the major table version indicates that the neighbor is not yet fully updated. The update interval for a neighbor in another AS is normally 30 seconds (the default value of the BGP advertisement timer).

In addition to the information about all sessions to all neighbors, the output also shows the amount of memory being used for the BGP data structures.

Practice

Q1) Which three steps must be taken before a BGP session is established? (Choose three.)

A) Neighbors must be manually configured as peers.

B) A TCP session between the two routers must be established.

C) The network topology database must be exchanged.

D) BGP Open messages must be exchanged and verified.

Q2) What three parameters are carried in a BGP Open message? (Choose three.)

A) BGP version number

B) local AS number

C) BGP hold timer

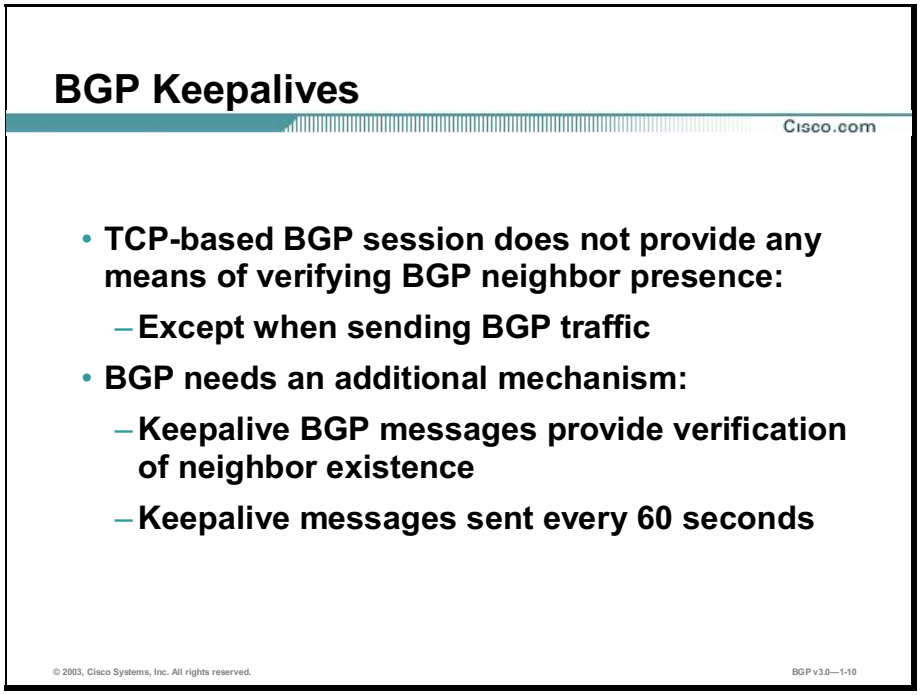
D) current neighbor state

E) network link-state database

 Copyright © 2003, Cisco Systems, Inc.

BGP Keepalives

This topic explains the role of the BGP keepalive in session establishment and maintenance.

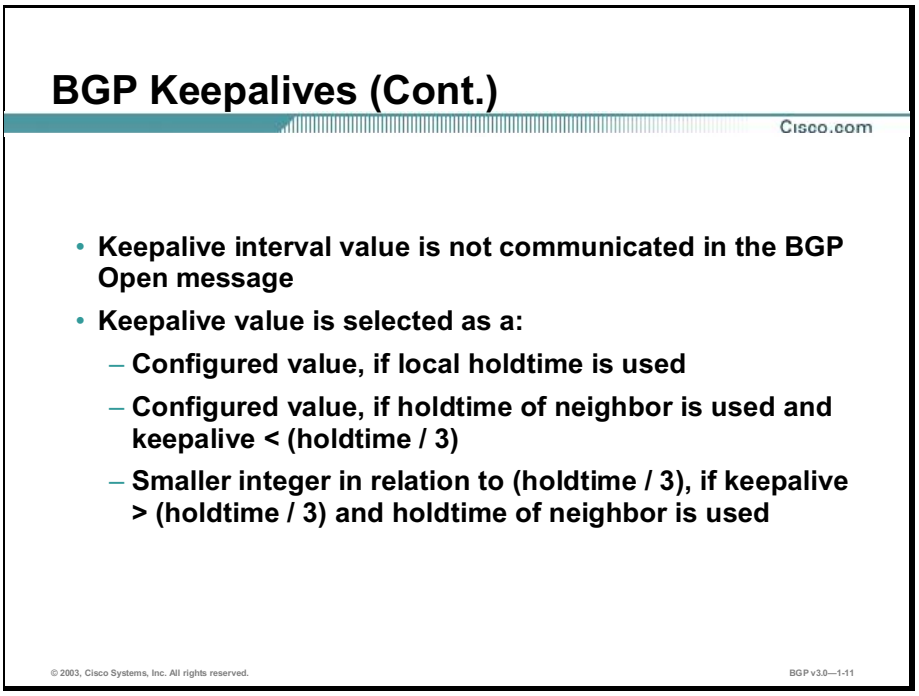

BGP Keepalives
Cisco.com
• TCP-based BGP session does not provide any means of verifying BGP neighbor presence:
– Except when sending BGP traffic
• BGP needs an additional mechanism:
– Keepalive BGP messages provide verification of neighbor existence
– Keepalive messages sent every 60 seconds
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-10

TCP-based BGP sessions do not provide any means of verifying the presence of a BGP neighbor. After BGP has established the TCP session, the only method of verifying neighbor presence is to actually send BGP traffic. BGP traffic is sent over the TCP session with acknowledgements, and is therefore reliable. Successfully sending BGP traffic confirms the existence of a BGP neighbor.

However, there are often long periods of time when no BGP traffic is sent between neighbors. During those periods, TCP implements no mechanism to check for the existence of the configured neighbor. BGP neighbors could therefore easily be disconnected during times of session inactivity. This situation would lead to incorrect routing information on the other side of the BGP session.

To avoid routing packets to a router that is no longer there, BGP needs an additional mechanism to make sure that a neighbor exists. BGP sends special keepalive messages during every keepalive interval to inform its peer of its presence. By default, this interval happens every 60 seconds. If no BGP traffic is received within the selected holdtime interval, the BGP router sends a BGP notification message to the inactive peer and tears down the BGP session between them. The default BGP holdtime value is 180 seconds.

When changing the default values of keepalive and holdtime intervals, you must take care not to configure too big a keepalive interval in comparison to the holdtime. Too big a difference could result in resetting of the BGP session after only one keepalive message has been missed, making a network unstable. The suggested ratio of keepalive to holdtime interval is 1:3.



As opposed to the holdtime interval, BGP peers do not communicate the keepalive interval in the Open message. The selection of a keepalive interval is therefore based on the selected holdtime value. The selected holdtime value used by both peers is the smaller one of both configured values.

The BGP process selects the keepalive interval value using following steps:

Step 1 If the locally configured value of holdtime is selected (being the lower of two), the peers will use the locally used keepalive interval.

Step 2 If the holdtime interval of the neighbor is selected, and the locally configured keepalive interval is less than a third of the holdtime interval, the peers will use the locally configured keepalive.

Step 3 If the holdtime interval of the neighbor is selected, and the locally configured keepalive interval is more than a third of the holdtime interval, the peers will use the smaller integer value in relation to (holdtime / 3).

Example

If the selected holdtime equals 17 seconds and the configured keepalive equals 10 seconds, the (holdtime / 3) rule will be used to select the keepalive value. Therefore, (17 / 3) = 5.67 and the keepalive value used by BGP will equal 5 seconds.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What is the purpose of the keepalive in BGP?

A) to provide the clock used for packet transmission

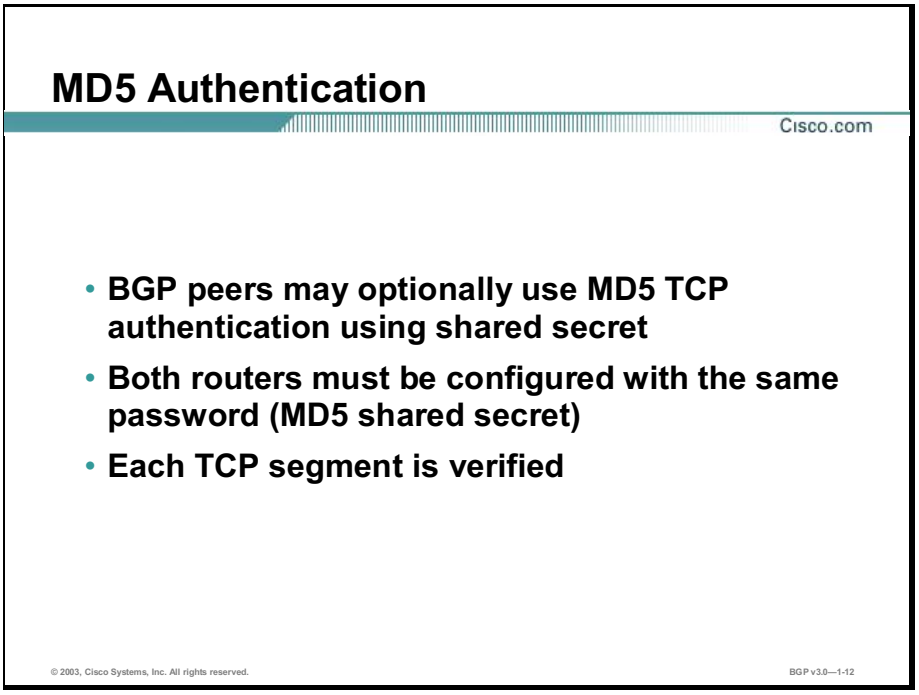
B) to synchronize the TCP neighbor sessions

C) to provide reliable transmission when you are using BGP between neighbors

D) to verify the presence of a configured BGP neighbor

MD5 Authentication

This topic describes how Message Digest 5 (MD5) authentication protects a BGP neighbor session.


MD5 Authentication
Cisco.com
• BGP peers may optionally use MD5 TCP authentication using shared secret
• Both routers must be configured with the same password (MD5 shared secret)
• Each TCP segment is verified
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-12

Authentication between BGP neighbors can be negotiated between BGP speaking routers using optional parameters in the Open message. If you are using MD5 authentication, every TCP segment on the BGP session will be transmitted to the configured neighbor along with a checksum. The checksum is calculated together with a secret known by the two routers using the MD5 algorithm. The common secret is never transmitted on the network. If the receiver, which is using the same common secret, calculates the same checksum from the TCP segment, then the receiver can be pretty sure that the information is transmitted from the correct source and the information has not been altered.

Authentication of BGP sessions is a vital tool to avoid denial-of-service (DoS) attacks.

Practice

Q1) What are the benefits of using MD5?

A) It encrypts routing updates between peers.

B) It provides for payload compression.

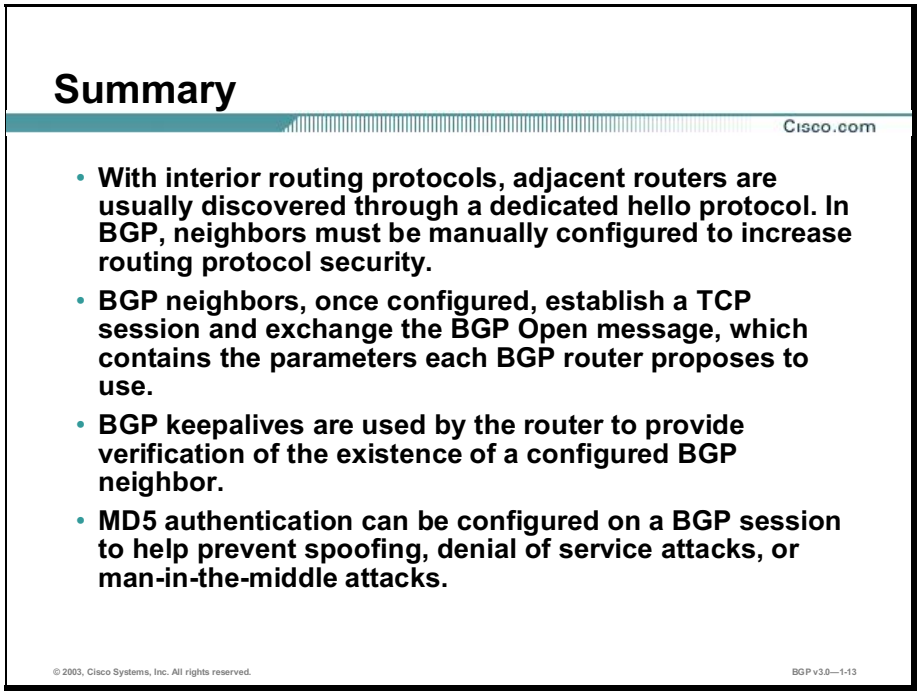
C) It ensures a reliable connection with retransmission of lost packets.

D) It provides peer authentication to prevent DoS attacks.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• With interior routing protocols, adjacent routers are usually discovered through a dedicated hello protocol. In BGP, neighbors must be manually configured to increase routing protocol security.
• BGP neighbors, once configured, establish a TCP session and exchange the BGP Open message, which contains the parameters each BGP router proposes to use.
• BGP keepalives are used by the router to provide verification of the existence of a configured BGP neighbor.
• MD5 authentication can be configured on a BGP session to help prevent spoofing, denial of service attacks, or man-in-the-middle attacks.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-13

Next Steps

After completing this lesson, go to:

- BGP Route Processing lesson

References

For additional information, refer to these resources:

- For more information on basic BGP, refer to “Border Gateway Protocol” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

- For more information on BGP neighbors and session establishment, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

Quiz: BGP Session Establishment

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe how BGP discovers neighbors
- Explain BGP session establishment procedures
- Explain the role of the BGP keepalive in session establishment and maintenance
- Describe how optional MD5 authentication can protect sessions between BGP peers

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What is indicated by a state of "Idle" in the output of the **show ip bgp summary** command?

A) The router is currently not attempting to establish a connection with a neighbor.

B) The connection to the configured neighbor has timed out.

C) The connection to a BGP neighbor has been established, and no errors have been received on the connection.

D) The connection to a BGP neighbor has been established, and no packets have been sent.

 Copyright © 2003, Cisco Systems, Inc.

Q2) What happens if two TCP connection attempts between configured BGP neighbors succeed?

A) Both connections will be terminated, and the neighbors will re-establish a neighbor relationship.

B) One connection will be maintained as primary and the other as backup.

C) One of the two connections will be torn down.

D) The router with the lower router-ID will determine if the second connection is torn down or used as a backup TCP connection.

Q3) Given the following BGP session states:

1. OpenConfirm
2. Established
3. Idle
4. OpenSent
5. Active

What is their order of progression during the creation of a successful neighbor session?

A) 5, 1, 4, 2, 3

B) 3, 4, 1, 5, 2

C) 5, 4, 1, 3, 2

D) 3, 5, 4, 1, 2

Q4) What does the field “TblVer” indicate in the output of the **show ip bgp summary** command?

A) the current version of BGP in use by the router

B) the number of route prefixes contained in the BGP update of the router

C) BGP messages received from that neighbor

D) the last version of the BGP database that was sent to that neighbor

Q5) What occurs when you use MD5 between two BGP neighbors?

A) Every packet is encrypted with MD5.

B) The IP header is encrypted using MD5.

C) An MD5 checksum is calculated and sent with each packet so that its source can be verified.

D) A username and password is embedded in an IP datagram that is matched to a username and password on the remote neighbor.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

BGP Route Processing

Overview

This lesson explains the details of processing IP routing information in Border Gateway Protocol (BGP). It describes how a router builds the BGP routing table, how BGP selects the best route, and how BGP routes are propagated to other BGP neighbors. The lesson also discusses how a router builds the IP routing table when it is using BGP.

Importance

Route processing is fundamental to the operation of BGP. Knowledge of the BGP route selection process, route propagation, and how the BGP and IP routing tables are built is key to properly configuring BGP and troubleshooting BGP routing issues.

Objectives

Upon completing this lesson, you will be able to:

- Explain BGP routing updates
- Describe how a router builds BGP tables
- Explain BGP route selection criteria
- Describe BGP route propagation
- Describe how a router builds an IP routing table when it is using BGP
- Describe how BGP advertises local networks
- Describe the role of automatic summarization in BGP route processing

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Successful completion of *Building Scalable Cisco Internetworks* (BSCI) or equivalent

Outline

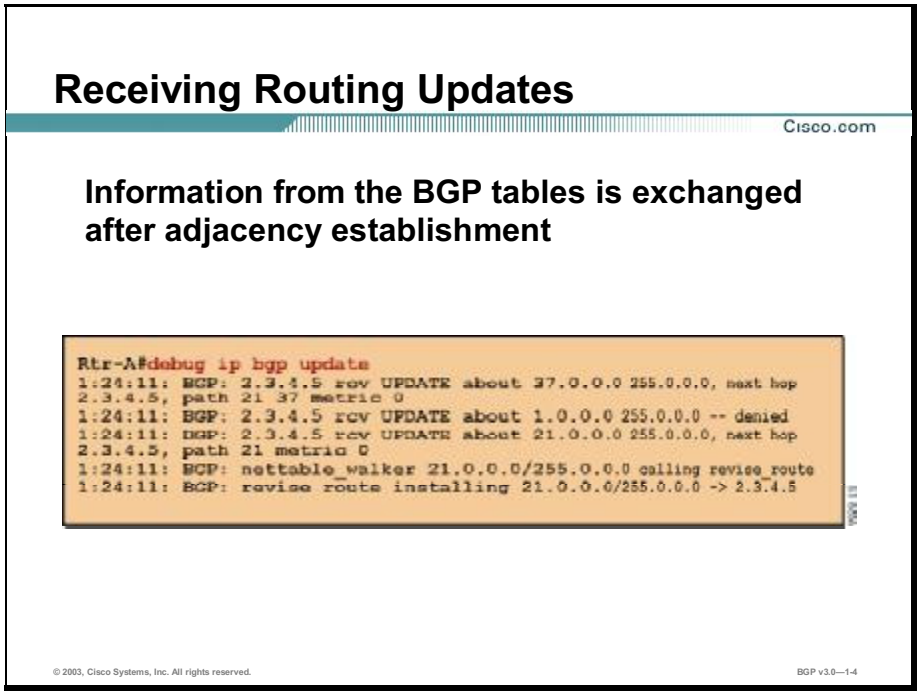
This lesson includes these topics:

- Overview
- Receiving Routing Updates
- Building the BGP Table
- BGP Route Selection Criteria
- BGP Route Propagation
- Building the IP Routing Table
- Advertising Local Networks
- Automatic Summarization
- Summary
- Assessment (Quiz): BGP Route Processing

 Copyright © 2003, Cisco Systems, Inc.

Receiving Routing Updates

This topic explains BGP routing updates.


Receiving Routing Updates
Cisco.com
Information from the BGP tables is exchanged after adjacency establishment
Rtr-A#debug ip bgp update
1:24:11: BGP: 2.3.4.5 rcv UPDATE about 37.0.0.0 255.0.0.0, next hop 2.3.4.5, path 21 37 metric 0
1:24:11: BGP: 2.3.4.5 rcv UPDATE about 1.0.0.0 255.0.0.0 -- denied
1:24:11: BGP: 2.3.4.5 rcv UPDATE about 21.0.0.0 255.0.0.0, next hop 2.3.4.5, path 21 metric 0
1:24:11: BGP: nettable_walker 21.0.0.0/255.0.0.0 calling revise_route
1:24:11: BGP: revise route installing 21.0.0.0/255.0.0.0 -> 2.3.4.5
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-4

After a BGP session is established, routing updates start to arrive. Each BGP routing update consists of one or more entries (routes). Each route is described according to the IP address and subnet mask along with any number of attributes. The next-hop, AS-path and origin attributes must always be present. Other BGP attributes are optionally present.

The debug output shows how information about network 37.0.0.0/8 is received from neighbor 2.3.4.5. The neighbor indicates that IP packets to destination IP addresses in network 37.0.0.0/8 can be forwarded to the next-hop address 2.3.4.5. The AS path 21 37 indicates that the final destination is in AS 37 but the packets have to pass through AS 21 in order to get there. The metric is the multi-exit discriminator (MED) value.

Network 21.0.0.0/8 also has the next-hop address of 2.3.4.5, but the AS path of 21 indicates that that network is inside the directly connected AS 21.

Network 1.0.0.0/8 is denied. The reason is not obvious from the debug output, but the network topology information indicates that network 1.0.0.0 is local to (inside) AS 123. AS 123 has advertised the network to AS 37, which propagates to AS 21 and returns back to AS 123. This information loop is detected by the content in the AS-path attribute. The receiving router detects its own AS number in the AS path and silently discards (denies) the route.

Practice

Q1) What three BGP attributes will always be present in received BGP routing updates? (Choose three.)

A) next-hop

B) local preference

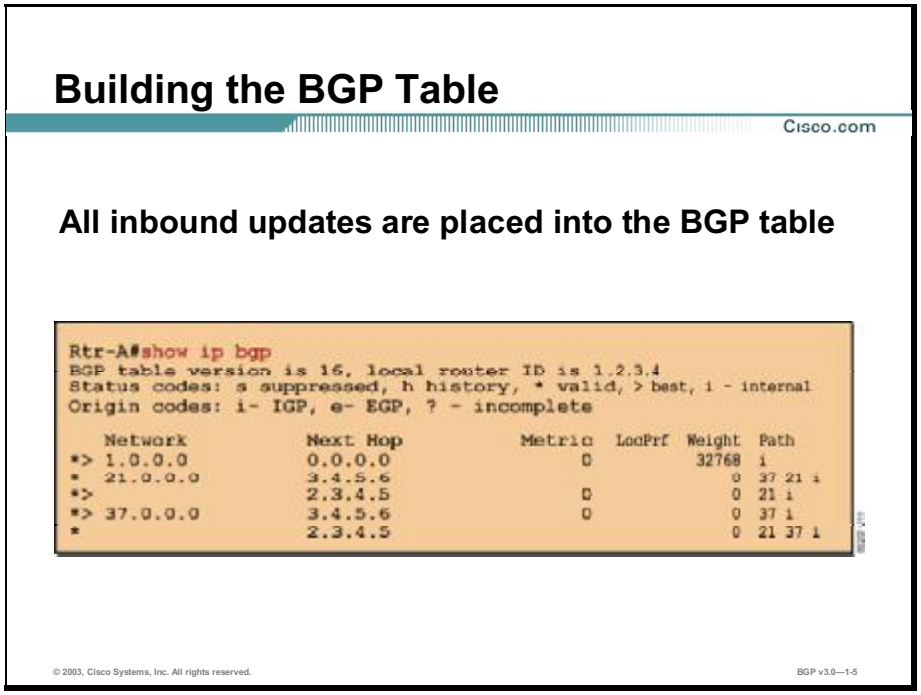
C) origin

D) AS-path

 Copyright © 2003, Cisco Systems, Inc.

Building the BGP Table

This topic explains how a router enabled for BGP builds a BGP route table.


Building the BGP Table
Cisco.com
All inbound updates are placed into the BGP table
Rtr-A#show ip bgp
BGP table version is 16, local router ID is 1.2.3.4
Status codes: s suppressed, h history, * valid, > best, i - internal
Origin codes: i- IGP, e- EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 1.0.0.0 0.0.0.0 0 32768 i
* 21.0.0.0 3.4.5.6 0 37 21 i
*> 2.3.4.5 0 0 21 i
*> 37.0.0.0 3.4.5.6 0 0 37 i
* 2.3.4.5 0 21 37 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-5

All routes received from a neighbor are saved in the router memory. Therefore, there is no need to retransmit or refresh any unchanged information.

When there is more than one way to reach a particular network, the local router selects one of those as the best. If that alternative is later lost because the neighboring router withdraws the route (or the neighboring router is no longer reachable), the remaining alternatives are still stored in memory and a new alternative is selected as the best without involving other BGP routers.

The **show ip bgp** router command gives an overview of all routing information received from all neighbors. The command will display basic information about each route on a single line. The output is sorted—different alternatives to reach the same network are displayed on consecutive lines. The network number is displayed only on the first lines indicating the same network. The network column is left blank on the consecutive lines indicating alternatives to reach the same network.

The router selects only one of the alternatives as the best path toward the destination. This alternative is indicated with the “>” sign.

Practice

Q1) Which incoming BGP updates are stored in the BGP table?

A) only the routes selected as the best

B) only the routes that are learned from external BGP peers

C) all routes received from BGP neighbors

D) all routes received from Interior Gateway Protocols

Q2) What does the router use in the output of the **show ip bgp** command to indicate the best route?

A) i

B) *

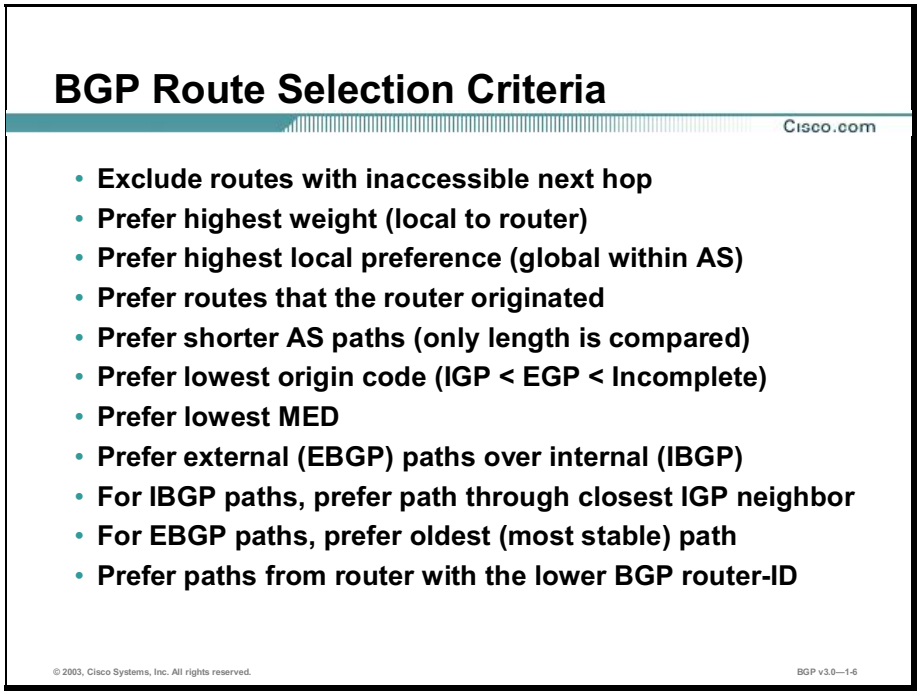
C) >

D) B

 Copyright © 2003, Cisco Systems, Inc.

BGP Route Selection Criteria

This topic explains the route selection process in BGP.


BGP Route Selection Criteria
Cisco.com
• Exclude routes with inaccessible next hop
• Prefer highest weight (local to router)
• Prefer highest local preference (global within AS)
• Prefer routes that the router originated
• Prefer shorter AS paths (only length is compared)
• Prefer lowest origin code (IGP < EGP < Incomplete)
• Prefer lowest MED
• Prefer external (EBGP) paths over internal (IBGP)
• For IBGP paths, prefer path through closest IGP neighbor
• For EBGP paths, prefer oldest (most stable) path
• Prefer paths from router with the lower BGP router-ID
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-6

When a router has more than one alternative route to reach the same IP subnet (network and mask), the router has to select one of them as best in its default mode of operation. In order to make this selection, the router uses the BGP attributes that are attached to the different updates. The selection criteria are checked in the order indicated in the following steps. The first of the checks that indicates a difference is used, and no further testing is done.

Step 1 The router checks if the next-hop attribute indicates an IP address that is reachable according to the current forwarding table. It is not necessary to have a direct connection to the next hop. It can very well be several router hops away and the route to it learned by the Interior Gateway Protocol (IGP). If the next hop is not reachable, the router does not consider the BGP route as a candidate to become selected as the best.

Step 2 The router prefers the route with the higher weight. The weight is not carried with the updates; it is a value assigned to the route by the local router and considered only within the router itself.

Step 3 If the local preference attributes are different, the route with the highest value is selected best.

Step 4 If one of the routes is injected into the BGP table by the local router, the local router prefers it to any routes received from other BGP routers.

Step 5 At this point, the lengths of the AS paths are compared (the content is not checked; only the number of autonomous systems in each AS path is counted). The route with the shortest length is selected.

Step 6 If the AS-path lengths are the same, the origin code is checked. BGP will prefer the path with the lowest origin type: IGP is lower than EGP, and EGP is lower than Incomplete.

Step 7 The router next compares MED values but only if it receives the updates from the same neighboring AS. Routes with a lower MED are preferred.

Step 8 At this point it is clear that the destination network is outside the local AS and that there is not much difference between the alternatives. Because the IP packets to the destination network must leave the AS, it is better that they do so as quickly as possible. If any of the alternatives are received from a BGP peer in another AS, that alternative is preferred.

Step 9 If the router receives all alternatives from peer routers in the local AS, each of them will indicate an exit point, and the closest exit is used. Distance to the exit point is calculated by comparing the IGP costs against the BGP next hops, as indicated in the forwarding table.

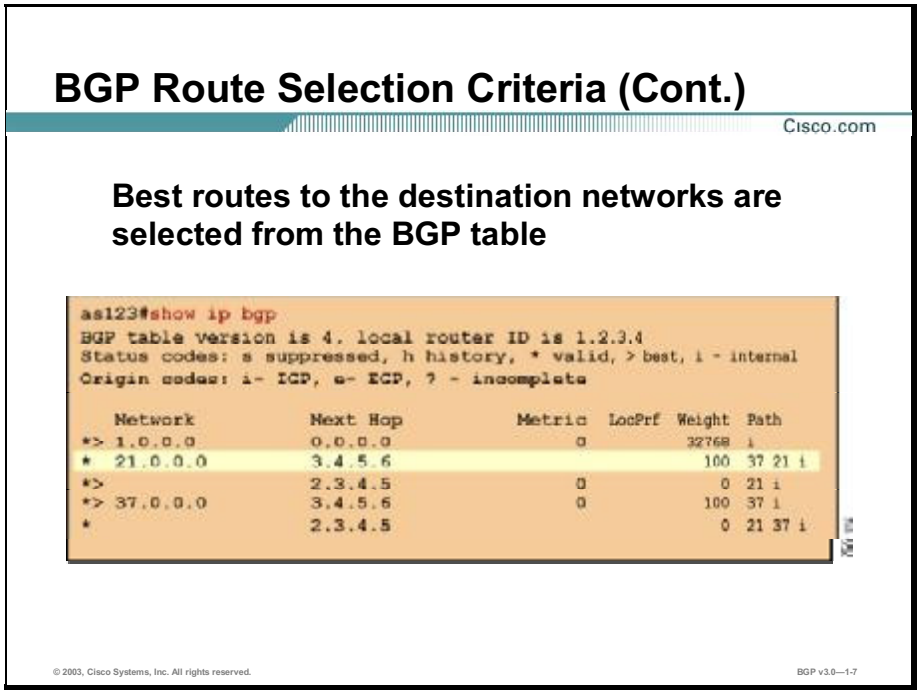
Step 10 If the router receives all alternatives from External Border Gateway Protocol (EBGP) neighbors, the most stable path (the oldest path) is preferred.

Step 11 If the router still cannot differentiate the routes, it nevertheless has to make a decision and select the best route. It checks the BGP sessions on which it received the updates and chooses the route received on the session where the peer router has the lower BGP router-ID.

The router makes the final test only after it has made all other checks and determined that all alternative routes are equally good.

 Copyright © 2003, Cisco Systems, Inc.

Example


BGP Route Selection Criteria (Cont.)
Cisco.com
Best routes to the destination networks are selected from the BGP table
as123#show ip bgp
BGP table version is 4, local router ID is 1.2.3.4
Status codes: s suppressed, h history, * valid, > best, i - internal
Origin codes: i- IGP, e- EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 1.0.0.0 0.0.0.0 0 32768 i
* 21.0.0.0 3.4.5.6 100 37 21 i
*> 2.3.4.5 0 0 21 i
*> 37.0.0.0 3.4.5.6 0 100 37 i
* 2.3.4.5 0 21 37 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-7

In this example, the router in AS 123 can reach network 21.0.0.0/8 via two paths; the first one is via neighbor 3.4.5.6 in AS 37 and then to AS 21, and the second path is straight to AS 21 through neighbor 2.3.4.5. In this example, the weight is set to 100 for the path via AS 37, and the other alternative path does not have the weight set. Thus, when checked against BGP path selection rules, the route via AS 37 is selected as the best because it has a higher weight attribute.

Likewise, network 37.0.0.0/8 is reached via AS 37 because the weight indicates it is the best route.

Practice

Q1) With no BGP attributes modified, which criterion will most likely determine route selection?

A) weight

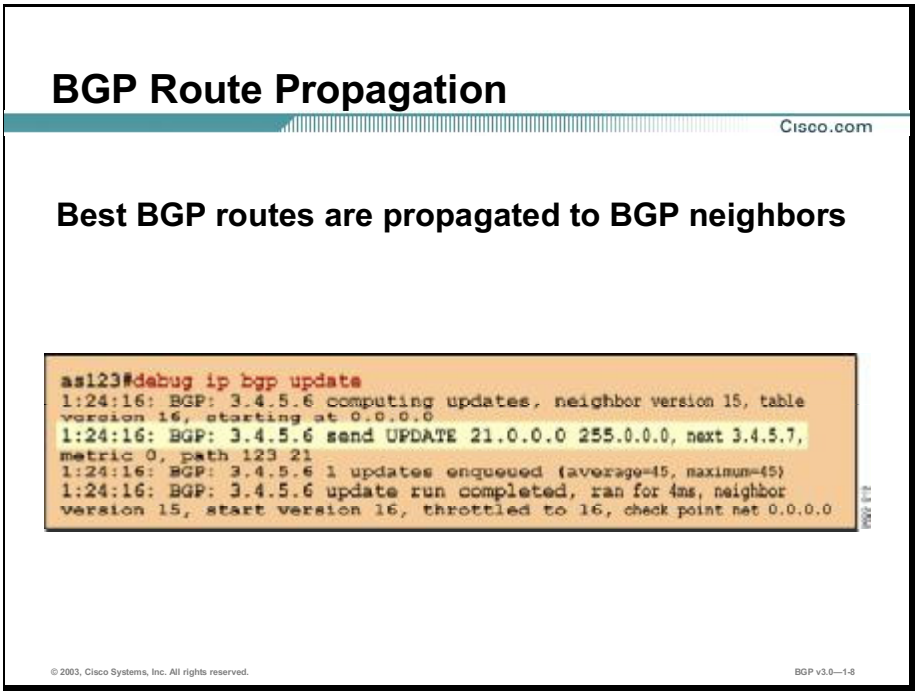
B) local preference

C) MED

D) AS-path length

BGP Route Propagation

This topic describes how a router propagates BGP routes to other BGP neighbors.


BGP Route Propagation
Cisco.com
Best BGP routes are propagated to BGP neighbors
as123#debug ip bgp update
1:24:16: BGP: 3.4.5.6 computing updates, neighbor version 15, table version 16, starting at 0.0.0.0
1:24:16: BGP: 3.4.5.6 send UPDATE 21.0.0.0 255.0.0.0, next 3.4.5.7, metric 0, path 123 21
1:24:16: BGP: 3.4.5.6 1 updates enqueued (average=45, maximum=45)
1:24:16: BGP: 3.4.5.6 update run completed, ran for 4ms, neighbor version 15, start version 16, throttled to 16, check point net 0.0.0.0
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-8

A local router propagates only the route selected as best to the neighbors. However, the router never sends a route back on the same BGP session upon which it was received. On the contrary, when it selects a neighbor as the best next hop, the router makes sure that the neighbor is not pointing back to the local router; it accomplishes this task by “poisoning” the route (marking the route unreachable) and sending a withdraw message to that neighbor.

The router conducts route poisoning to avoid a potential routing loop problem in which a neighbor router selected as the best next hop might rely on the local router as the best next hop.

The process where routing information is not sent back to the source of information is called split horizon.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Which BGP routes are announced to other BGP neighbors?

A) All learned routes are sent to all BGP neighbors.

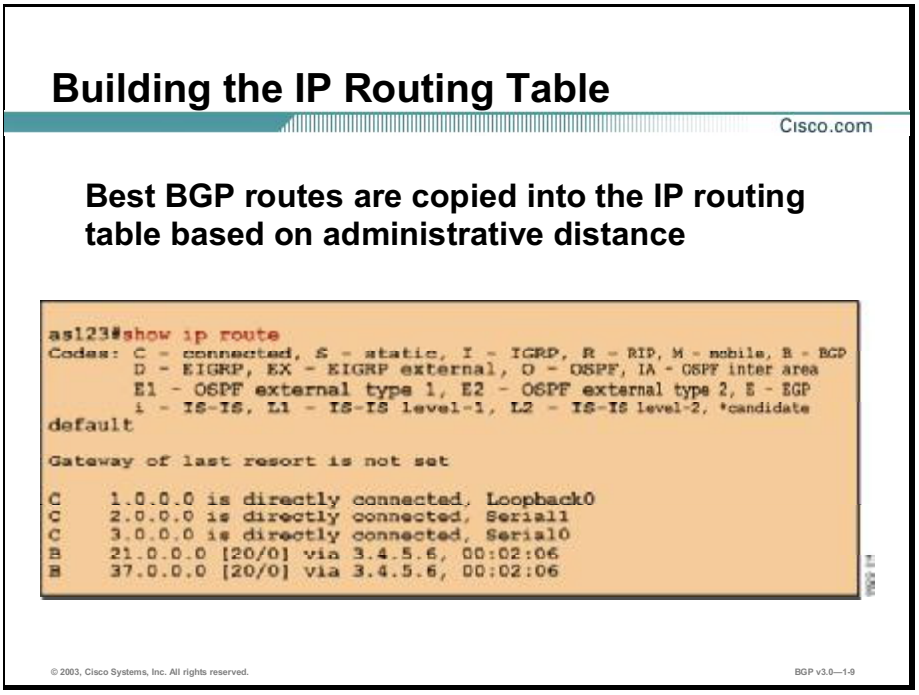
B) Only the routes learned from IGPs are sent to other neighbors.

C) Only the routes selected as the best are announced to other neighbors.

D) All routes learned from EBGP neighbors are sent to all other BGP peers.

Building the IP Routing Table

This topic explains the process of building an IP routing table from the BGP table and from other sources of routing information such as IGPs.


Building the IP Routing Table
Cisco.com
Best BGP routes are copied into the IP routing table based on administrative distance
as123#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, *candidate
default
Gateway of last resort is not set
C 1.0.0.0 is directly connected, Loopback0
C 2.0.0.0 is directly connected, Serial1
C 3.0.0.0 is directly connected, Serial0
B 21.0.0.0 [20/0] via 3.4.5.6, 00:02:06
B 37.0.0.0 [20/0] via 3.4.5.6, 00:02:06
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-9

The route in the BGP table that BGP selects as the best is a candidate for installation in the IP forwarding, or routing, table.

Before a route can be installed, the router has to check if there is any other routing protocol that has information about the same subnet (network and mask). If the subnet is known via different sources, the router uses the administrative distance (AD) to determine which source to use. The router will install the route with the lowest AD.

The output from the **show ip route** command indicates which routes in the forwarding table were installed using the BGP information. Those routes are denoted with the letter “B.” The AD is shown in the command output as the first number within the brackets.

In this example, both networks 21.0.0.0/8 and 37.0.0.0/8 are reachable via 3.4.5.6. After the router has installed the routes in the forwarding table, user data traffic starts to be forwarded.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What routing characteristic does a router use to select the best route when multiple sources exist for the same route?

A) cost

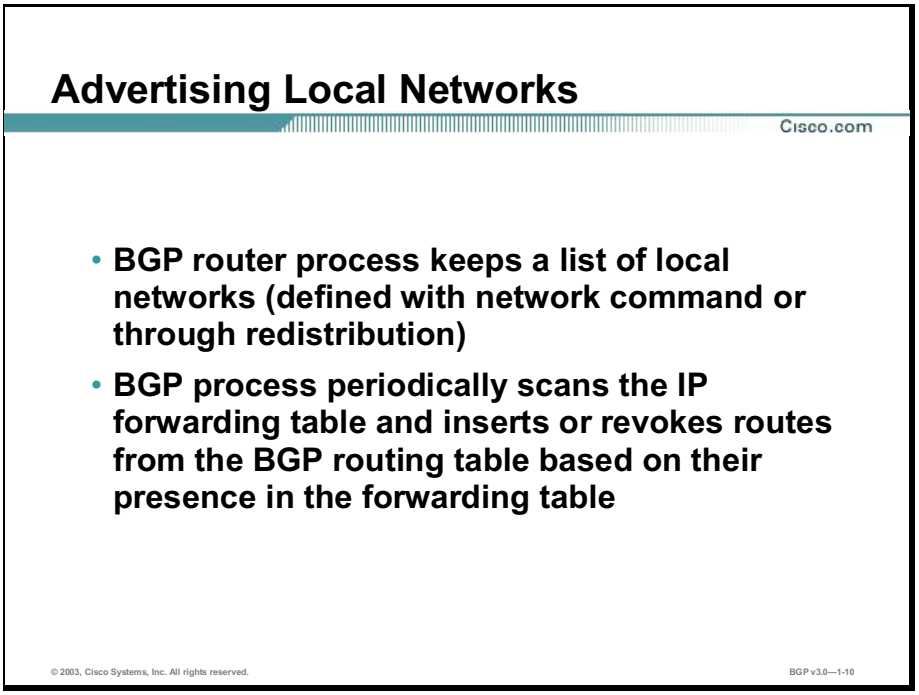
B) AD

C) weight

D) metric

Advertising Local Networks

This topic discusses how BGP advertises local networks.


Advertising Local Networks
Cisco.com
• BGP router process keeps a list of local networks (defined with network command or through redistribution)
• BGP process periodically scans the IP forwarding table and inserts or revokes routes from the BGP routing table based on their presence in the forwarding table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-10

The BGP routing process can inject new routes in the BGP table. A router will propagate newly injected routes to neighboring BGP peers if it selects them as best, giving neighboring autonomous systems information about networks that are reachable in the local AS. This process is called advertising, originating, or announcing, local routes.

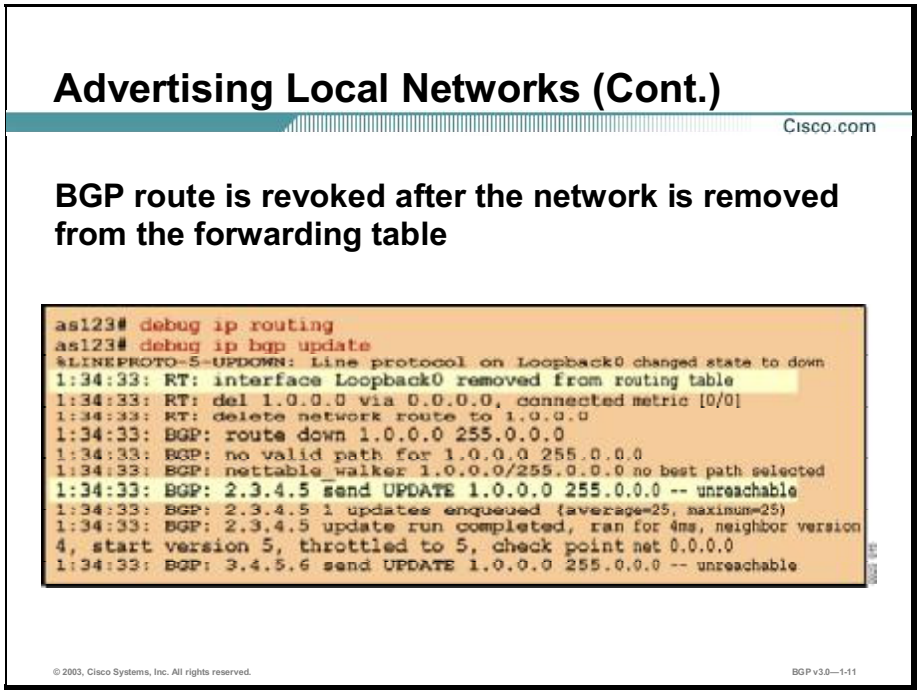
The BGP process can inject local routes in two different ways:

- A list of networks is configured on the router under the BGP router process using the **network** configuration command. The networks listed are candidates for being injected. Networks are injected only if they appear in the forwarding table. In the case where the IGP used within the AS finds a valid path to them, the routes will be in the forwarding table.

- Redistribution of routes learned by another routing protocol. The IGP used with the AS can also act as a source of routing information about local networks.

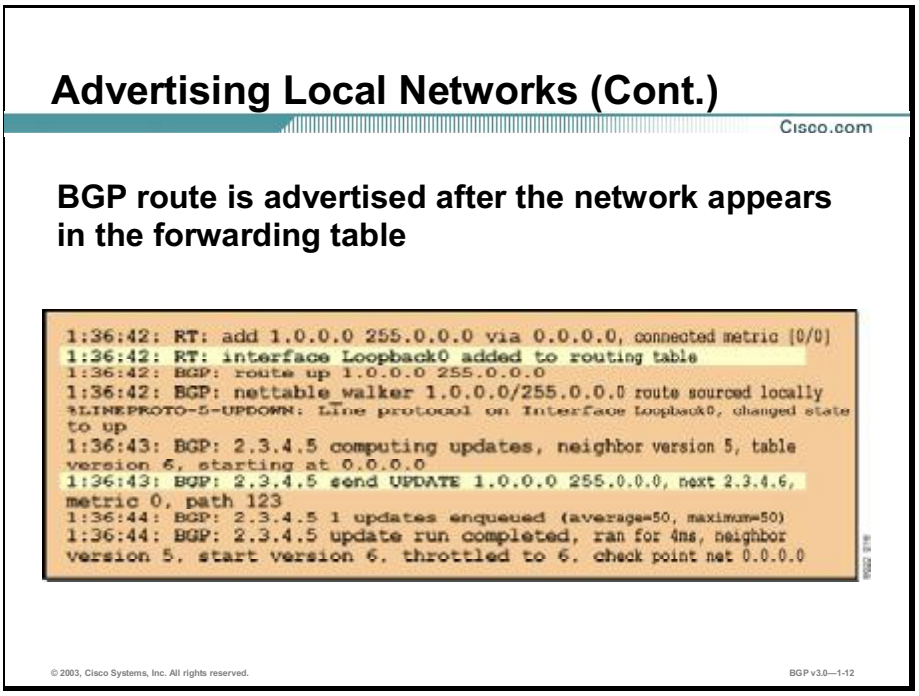
 Copyright © 2003, Cisco Systems, Inc.

Example


Advertising Local Networks (Cont.)
Cisco.com
BGP route is revoked after the network is removed from the forwarding table
as123# debug ip routing
as123# debug ip bgp update
%LINEPROTO-5-UPDOWN: Line protocol on Loopback0 changed state to down
1:34:33: RT: interface Loopback0 removed from routing table
1:34:33: RT: del 1.0.0.0 via 0.0.0.0, connected metric [0/0]
1:34:33: RT: delete network route to 1.0.0.0
1:34:33: BGP: route down 1.0.0.0 255.0.0.0
1:34:33: BGP: no valid path for 1.0.0.0 255.0.0.0
1:34:33: BGP: nettable_walker 1.0.0.0/255.0.0.0 no best path selected
1:34:33: BGP: 2.3.4.5 send UPDATE 1.0.0.0 255.0.0.0 -- unreachable
1:34:33: BGP: 2.3.4.5 1 updates enqueued (average=25, maximum=25)
1:34:33: BGP: 2.3.4.5 update run completed, ran for 4ms, neighbor version
4, start version 5, throttled to 5, check point net 0.0.0.0
1:34:33: BGP: 3.4.5.6 send UPDATE 1.0.0.0 255.0.0.0 -- unreachable
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-11

In this example, network 1.0.0.0/8 is directly connected to interface loopback 0. The route to 1.0.0.0/8 has been previously installed in the BGP table because it was listed with a **network** statement and it was in the forwarding table as directly connected. When the loopback 0 interface goes down, the router removes the directly connected route from its routing table. Since the route no longer exists in the routing table, it must also be removed from the BGP table.

Because there has been a change in the BGP table, the BGP neighbors must be informed. The router sends a BGP update message to both neighbors indicating that network 1.0.0.0/8 is now unreachable.


Advertising Local Networks (Cont.)
Cisco.com
BGP route is advertised after the network appears in the forwarding table
1:36:42: RT: add 1.0.0.0 255.0.0.0 via 0.0.0.0, connected metric [0/0]
1:36:42: RT: interface Loopback0 added to routing table
1:36:42: BGP: route up 1.0.0.0 255.0.0.0
1:36:42: BGP: nettable_walker 1.0.0.0/255.0.0.0 route sourced locally
%LINEPROTO-5-UPDOWN: Line protocol on Interface Loopback0, changed state to up
1:36:43: BGP: 2.3.4.5 computing updates, neighbor version 5, table version 6, starting at 0.0.0.0
1:36:43: BGP: 2.3.4.5 send UPDATE 1.0.0.0 255.0.0.0, next 2.3.4.6, metric 0, path 123
1:36:44: BGP: 2.3.4.5 1 updates enqueued (average=50, maximum=50)
1:36:44: BGP: 2.3.4.5 update run completed, ran for 4ms, neighbor version 5, start version 6, throttled to 6, check point net 0.0.0.0
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-12

In this example, network 1.0.0.0/8 is listed with a **network** statement in the BGP process. However, the network was not in the router's forwarding table so it was not injected into its BGP table.

Now the loopback 0 interface comes back up again. This reappearance means that the network 1.0.0.0/8 is now in the forwarding table as a directly connected route. As a result, the router will once again inject the 1.0.0.0/8 network into its BGP table, and subsequently update its configured neighbors.

Practice

Q1) What are two methods for advertising local routes in BGP? (Choose two.)

A) announcing networks using the **network** command

B) sending and receiving routing updates using the **neighbor** command

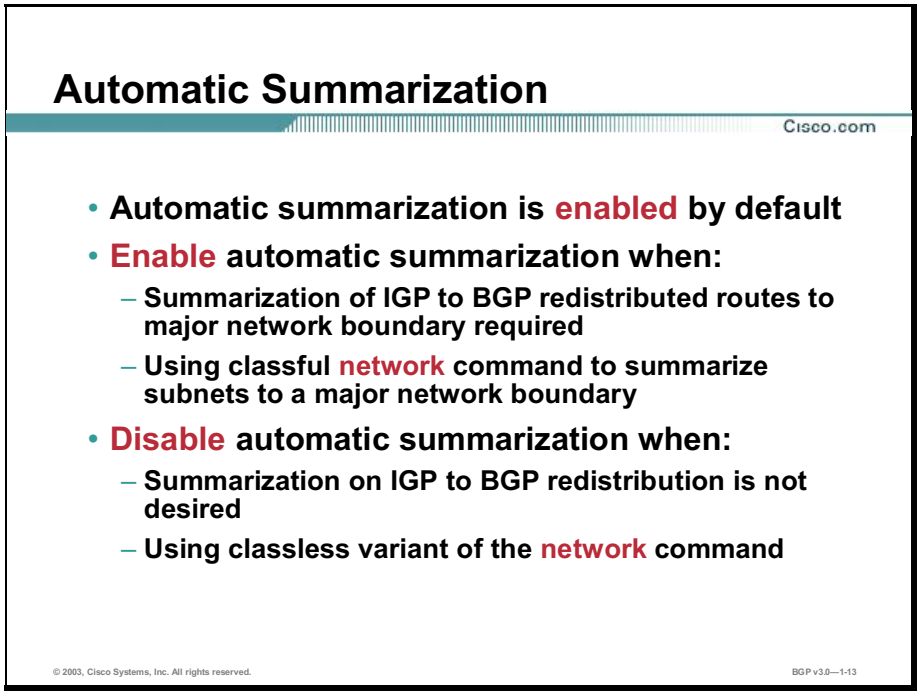
C) redistribution of routes learned by another routing protocol

D) automatic redistribution of local routes to BGP peers

 Copyright © 2003, Cisco Systems, Inc.

Automatic Summarization

This topic describes the role of automatic summarization in BGP route processing.


Automatic Summarization
Cisco.com
• Automatic summarization is enabled by default
• Enable automatic summarization when:
– Summarization of IGP to BGP redistributed routes to major network boundary required
– Using classful network command to summarize subnets to a major network boundary
• Disable automatic summarization when:
– Summarization on IGP to BGP redistribution is not desired
– Using classless variant of the network command
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-13

When a BGP router is configured to locally announce routes into BGP, the behavior of the **network** command varies depending on whether automatic summarization is enabled or disabled. When automatic summarization is enabled, BGP summarizes the locally originated BGP networks (network x.x.x.x) to their classful boundaries. Automatic summarization is enabled by default in BGP.

When a subnet exists in the routing table and the following three conditions are satisfied, then any subnet (component route) of that classful network in the local routing table will prompt BGP to install the classful network into the BGP table:

- A classful **network** statement for a network that exists in the routing table.
- A classful mask has been configured on that **network** statement.
- Automatic summarization is enabled.

When automatic summarization is disabled, the routes introduced locally into the BGP table are not summarized to their classful boundaries.

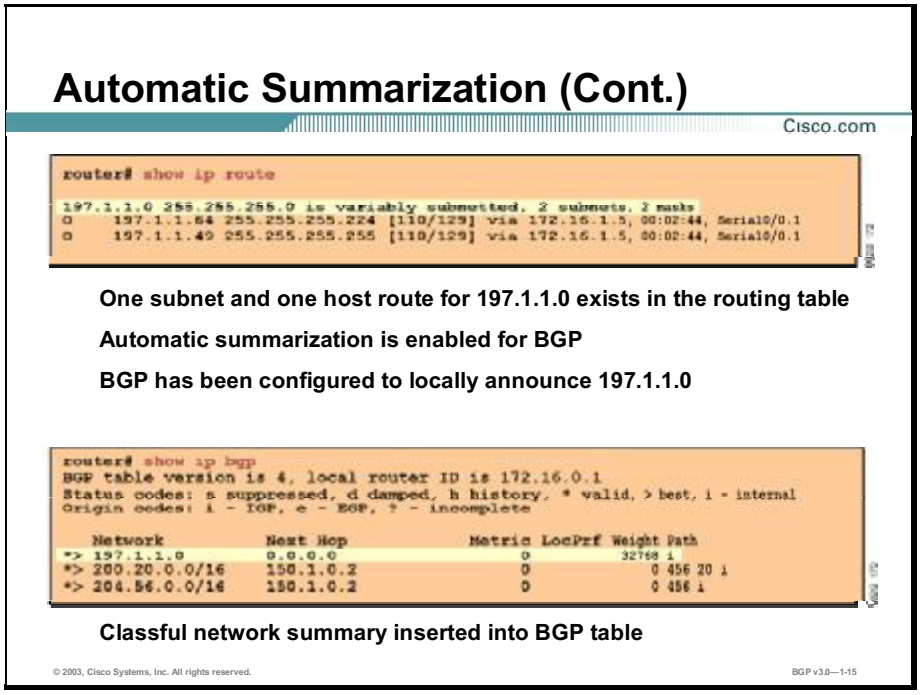
The behavior of the redistribution procedure in BGP is also influenced by the configuration of automatic summarization on the router. When enabled, all redistributed subnets will be summarized to their classful boundaries in the BGP table. When disabled, all redistributed subnets will be present in their original form in the BGP table.

Enable automatic summarization in BGP when the summarization of subnets to their classful boundaries will not introduce flawed information into the BGP table. In other words, leave automatic summarization enabled only when you are using a fully assigned classful network matching the network summarized in BGP.

Whenever possible, use the classless variant of the **network** command, specifying the subnet mask length of the network. When redistributing networks into BGP, the preferred method is to disable automatic summarization. Disabling automatic summarization ensures that correct information is inserted into the router's BGP table.

 Copyright © 2003, Cisco Systems, Inc.

Example


Automatic Summarization (Cont.)
Cisco.com
router# show ip route
197.1.1.0 255.255.255.0 is variably subnetted, 2 subnets, 2 masks
O 197.1.1.64 255.255.255.224 [110/129] via 172.16.1.5, 00:02:44, Serial0/0.1
O 197.1.1.49 255.255.255.255 [110/129] via 172.16.1.5, 00:02:44, Serial0/0.1
One subnet and one host route for 197.1.1.0 exists in the routing table
Automatic summarization is enabled for BGP
BGP has been configured to locally announce 197.1.1.0
router# show ip bgp
BGP table version is 4, local router ID is 172.16.0.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 197.1.1.0 0.0.0.0 0 32768 i
*> 200.20.0.0/16 150.1.0.2 0 0 456 20 i
*> 204.56.0.0/16 150.1.0.2 0 0 456 i
Classful network summary inserted into BGP table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-15

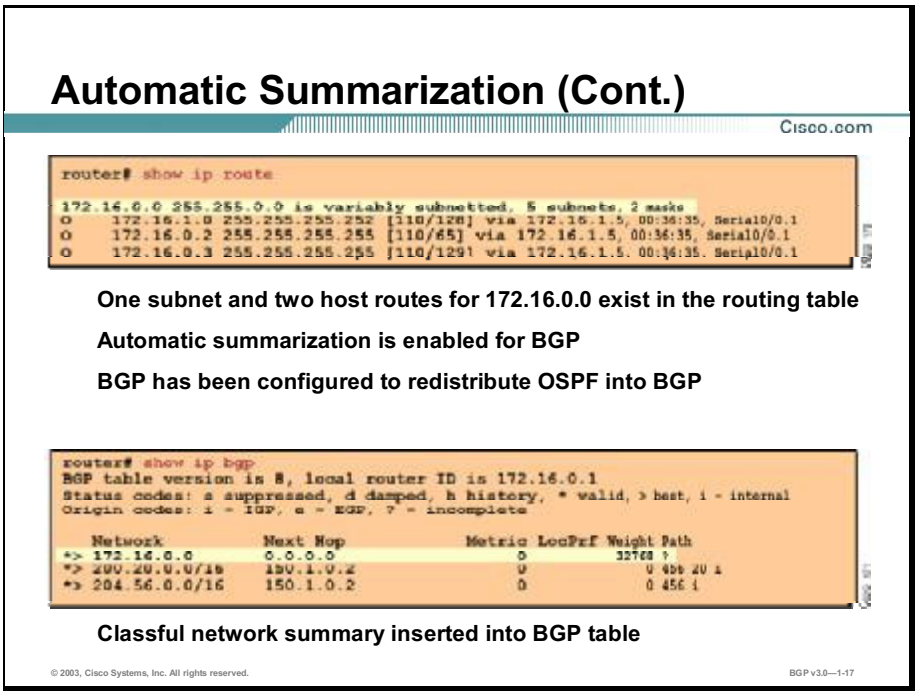
When you are inserting networks into the BGP table with the classful **network** command and automatic summarization disabled, no insertion into the BGP table will occur unless an exact match exists in the IP routing table (meaning a classful network has to be present in the IP routing table).

When automatic summarization is enabled, the major **network** command will summarize all subnets in the IP routing table to their major network boundary.

In this example, one subnet and one host route of the major class C network 197.1.1.0/24 (197.1.1.64/27 and 197.1.1.49/32) exists in the routing table. There is a classful **network** command, and automatic summarization is enabled for BGP. This setup results in the insertion of a classful network summary into the BGP table, instead of separate subnets.

Subnet 197.1.1.64/27 and host route 197.1.1.49/32 were summarized during insertion into the BGP table to the classful network 197.1.1.0/24. This action occurred because a classful **network** command and automatic summarization were configured on the router. If automatic summarization were disabled, no insertion into the BGP table would occur at all.

The locally sourced summary has all the attributes of a locally sourced BGP route (next hop=0.0.0.0, weight=32768, empty AS-path list), and IGP origin (being sourced with the **network** command).


Automatic Summarization (Cont.)
Cisco.com
router# show ip route
172.16.0.0 255.255.0.0 is variably subnetted, 5 subnets, 2 masks
O 172.16.1.0 255.255.255.252 [110/128] via 172.16.1.5, 00:36:35, Serial0/0.1
O 172.16.0.2 255.255.255.255 [110/65] via 172.16.1.5, 00:36:35, Serial0/0.1
O 172.16.0.3 255.255.255.255 [110/129] via 172.16.1.5, 00:36:35, Serial0/0.1
One subnet and two host routes for 172.16.0.0 exist in the routing table
Automatic summarization is enabled for BGP
BGP has been configured to redistribute OSPF into BGP
router# show ip bgp
BGP table version is 8, local router ID is 172.16.0.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 172.16.0.0 0.0.0.0 0 32768 ?
*> 200.20.0.0/16 150.1.0.2 0 0 456 20 i
*> 204.56.0.0/16 150.1.0.2 0 0 456 i
Classful network summary inserted into BGP table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-17

In this example, automatic summarization is enabled, resulting in the summarization of redistributed subnets to their classful boundaries. Subnet 172.16.1.0/30 and the two host routes 172.16.0.2/32 and 172.16.0.3/32 will be summarized into the single class B network 172.16.0.0/16. The network 172.16.0.0/16 is a locally sourced summary with all the attributes of a locally sourced BGP route (next hop=0.0.0.0, weight=32768, empty AS-path list). The origin of the route is marked as incomplete as the route is sourced through redistribution.

If automatic summarization were disabled, more specific routes would be present in the BGP table instead of the summary prefix 172.16.0.0/16.

Practice

Q1) Which combination of commands would result in inserting only more specific routes into the BGP table?

A) classful **network** command with automatic summarization enabled

B) redistribution into BGP with automatic summarization enabled

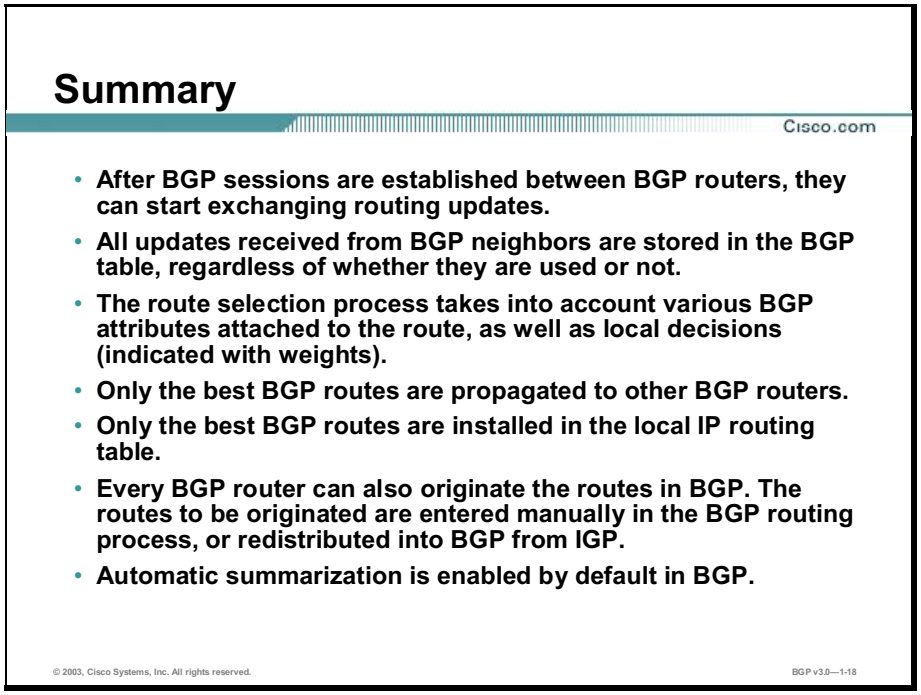
C) classful **network** command with automatic summarization disabled

D) redistribution into BGP with automatic summarization disabled

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• After BGP sessions are established between BGP routers, they can start exchanging routing updates.
• All updates received from BGP neighbors are stored in the BGP table, regardless of whether they are used or not.
• The route selection process takes into account various BGP attributes attached to the route, as well as local decisions (indicated with weights).
• Only the best BGP routes are propagated to other BGP routers.
• Only the best BGP routes are installed in the local IP routing table.
• Every BGP router can also originate the routes in BGP. The routes to be originated are entered manually in the BGP routing process, or redistributed into BGP from IGP.
• Automatic summarization is enabled by default in BGP.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-18

Next Steps

After completing this lesson, go to:

- Basic BGP Configuration lesson

References

For additional information, refer to these resources:

- For more information on BGP route processing, refer to “Border Gateway Protocol” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm
- For more information on BGP neighbors and session establishment, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

Quiz: BGP Route Processing

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Explain BGP routing updates
- Describe how a router builds BGP tables
- Explain BGP route selection criteria
- Describe BGP route propagation
- Describe how a router builds an IP routing table when it is using BGP
- Describe how BGP advertises local networks
- Describe the role of automatic summarization in BGP route processing

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What does a router running BGP do with a BGP update containing its own AS path?

A) The router checks to see if the information contained in the update is better than its current information. If it is, it will update its BGP table.

B) The router accepts the route update.

C) The router silently discards (denies) the route.

D) The router will return an error to the router sending the update.

 Copyright © 2003, Cisco Systems, Inc.

Q2) How many alternate paths to a single destination will a BGP router maintain in the BGP table?

A) The router will maintain only the best path to the destination.

B) The router will maintain two paths: the best path and a backup route.

C) The BGP table will hold up to four routes by default and a maximum of six configurable routes.

D) The BGP table will store all valid, advertised routes to the destination in the BGP table.

Q3) What are two ways in which local networks are advertised into the BGP routing protocol process? (Choose two.)

A) automatically, after a BGP neighbor session is established

B) manually, with the **network** command

C) through redistribution into the BGP process

D) by advertising them to the BGP table on the router after CDP discovers connected networks

Q4) What are two situations when it is appropriate to disable automatic summarization in BGP? (Choose two.)

A) when BGP neighbors are not configured to advertise aggregate routes to upstream providers

B) when the classless variant of the **network** command is used

C) when you are using a classless IGP in the AS

D) when the effects of automatic summarization of IGP-to-BGP redistribution are not desired

Q5) What is the AD of BGP routes in the IP routing table learned from BGP neighbors in a different AS?

A) 1

B) 20

C) 90

D) 120

Copyright © 2003, Cisco Systems, Inc.

Q6) What three BGP attributes are displayed for each route in the BGP table when you are using the **show ip bgp** command? (Choose three.)

A) weight

B) communities

C) origin

D) AS-path

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Basic BGP Configuration

Overview

This lesson introduces the Cisco IOS® commands required to configure a router for basic Border Gateway Protocol (BGP) operation. Included are the commands use to enable the BGP routing protocol process, establish neighbors, and advertise local routes. This lesson concludes with basic commands that network administrators can use to monitor the BGP configuration.

Importance

Basic BGP configuration is critical to any successful BGP implementation. Network administrators use the Cisco IOS commands included in this lesson in all BGP implementations. Thorough knowledge of the commands included within this lesson is therefore crucial to ensuring a successful implementation using BGP.

Objectives

Upon completing this lesson, you will be able to:

- Identify the Cisco IOS commands required to configure the BGP routing process
- Identify the Cisco IOS commands required to configure BGP neighbors
- Identify the Cisco IOS commands required to configure basic timers used in BGP
- Identify the Cisco IOS commands required to configure MD5 authentication for BGP
- Identify the commands required to announce local networks in BGP
- Describe BGP route redistribution and identify the commands required to configure BGP route redistribution

- Describe the classless behavior of BGP and identify the Cisco IOS commands required to configure BGP for classless operation
- Describe BGP route aggregation and identify the Cisco IOS commands required to configure basic BGP route aggregation
- Determine when BGP route aggregation is not appropriate in multihomed topologies

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Successful completion of *Building Scalable Cisco Internetworks* (BSCI) or equivalent

Outline

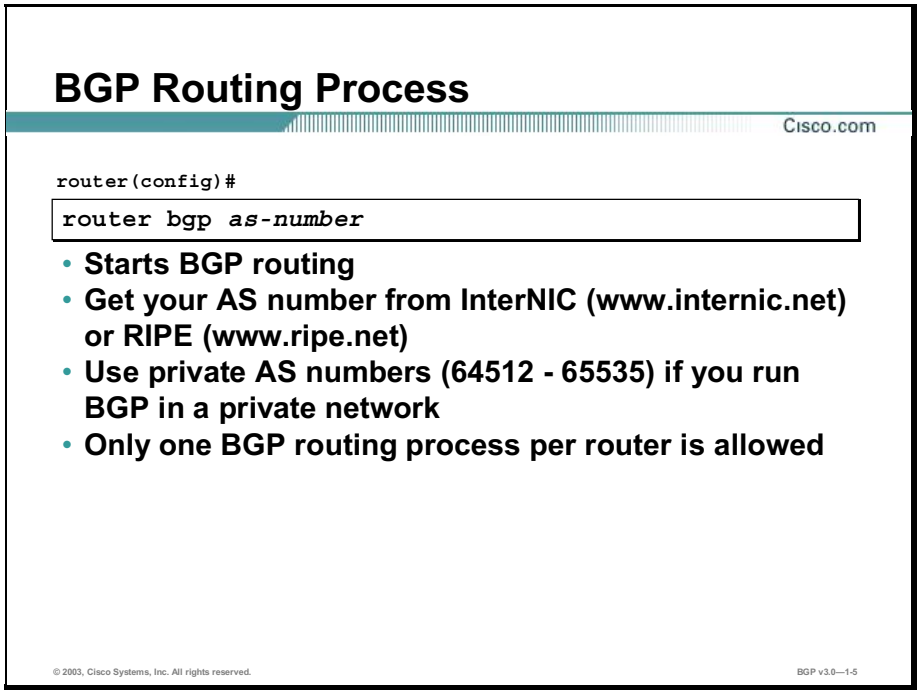
This lesson includes these topics:

- Overview
- BGP Routing Process
- Configuring External Neighbors
- Configuring BGP Timers
- Configuring MD5 Authentication
- Announcing Networks in BGP
- Redistributing Routes into BGP
- Configuring Classless BGP
- Aggregating BGP Networks
- Multihomed Customer Problem
- Summary
- Assessment (Quiz): Basic BGP Configuration

 Copyright © 2003, Cisco Systems, Inc.

BGP Routing Process

This topic identifies the commands required to initially configure the BGP routing process on a Cisco IOS router.


BGP Routing Process
Cisco.com
router(config)#
router bgp as-number
• Starts BGP routing
• Get your AS number from InterNIC (www.internic.net) or RIPE (www.ripe.net)
• Use private AS numbers (64512 - 65535) if you run BGP in a private network
• Only one BGP routing process per router is allowed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-5

router bgp

To configure the BGP routing process, use the **router bgp** global configuration command.

router bgp *autonomous-system*

To remove a routing process, use the **no** form of this command.

no router bgp *autonomous-system*

Syntax Description

autonomous-system	Number of an autonomous system (AS) that identifies the router to other BGP routers and tags the routing information passed along

This command starts the BGP routing process in the router. There can be, at most, one BGP process in a router. It must be assigned the local AS number.

The AS number is a 16-bit unsigned integer number. It must uniquely identify the AS among all routers exchanging BGP routing information, either directly or indirectly. This means that the AS numbers are required to be unique when BGP information is exchanged with the Internet.

The AS number can be a public AS number (ranging from 1 to 64511) assigned by an Internet registry (InterNIC: www.internic.com or RIPE: www.ripe.com), or a private AS number (ranging from 64512 to 65535). Private AS numbers will never be propagated onto the public Internet.

Practice

Q1) How many BGP processes can be active in a router?

A) 1

B) 16

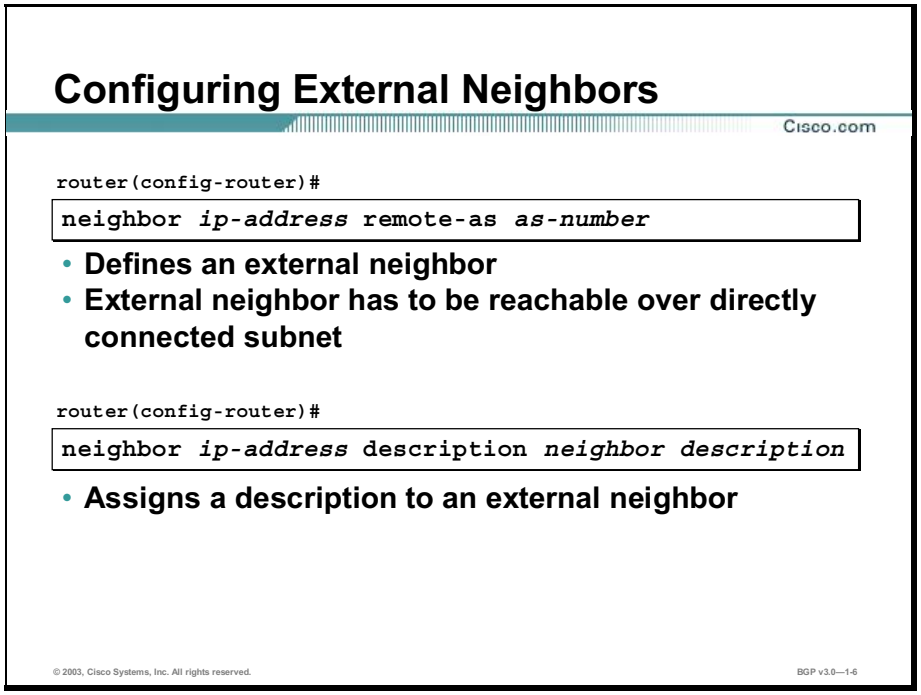
C) 256

D) depends on the configured memory in the router

 Copyright © 2003, Cisco Systems, Inc.

Configuring External Neighbors

This topic lists the commands required to configure external BGP neighbors on a Cisco router.


Configuring External Neighbors
Cisco.com
router(config-router)#
neighbor ip-address remote-as as-number
• Defines an external neighbor
• External neighbor has to be reachable over directly connected subnet
router(config-router)#
neighbor ip-address description neighbor description
• Assigns a description to an external neighbor
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-6

BGP does not automatically discover neighbors. They have to be explicitly configured. The local router will try to connect to the indicated IP address and also accept incoming connection attempts from the indicated IP address.

The first attribute that you must configure with a new neighbor is the remote-AS number in which the neighbor is taking part. When the TCP session is established between BGP routers, the configured remote-AS will be verified by each router with the exchange of BGP Open messages.

You may optionally configure other attributes with the neighbor. Do this on successive configuration lines, referring to the same neighbor IP address but indicating different attributes. In this figure, a description (text string) can be entered, describing the neighbor with the **neighbor description** command.

neighbor remote-as

To add an entry to the BGP neighbor table, use the **neighbor remote-as** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **remote-as** *number*

To remove an entry from the table, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **remote-as** *number*

Syntax Description

ip-address	IP address of neighbor
peer-group-name	Name of a BGP peer group
number	AS to which the neighbor belongs

neighbor description

To associate a description with a neighbor, use the **neighbor description** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **description** *text*

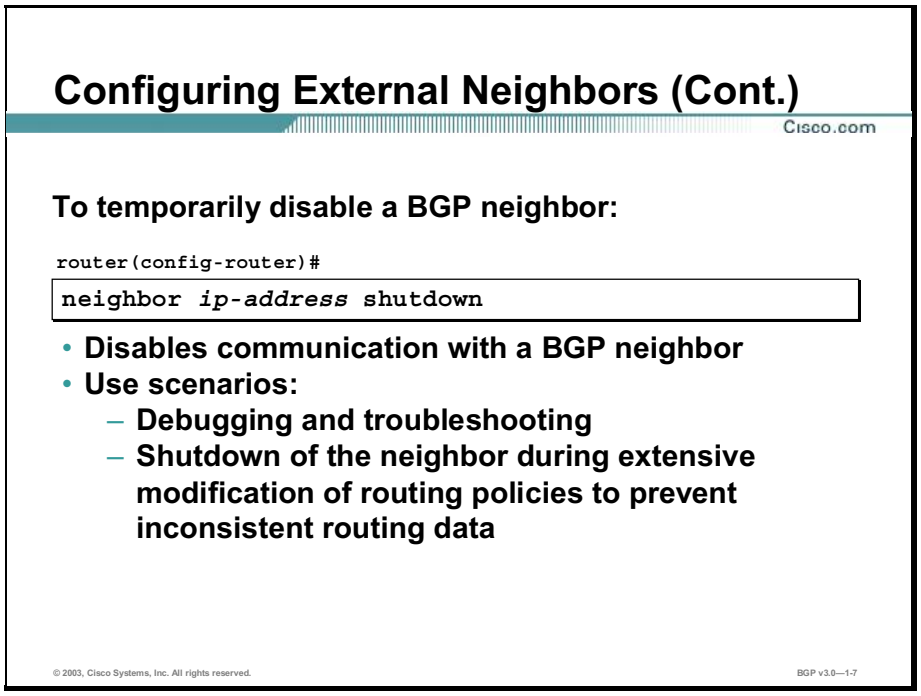
To remove the description, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **description** [*text*]

Syntax Description

ip-address	IP address of neighbor
peer-group-name	Name of a BGP peer group
text	Text (up to 80 characters) that describes the neighbor

 Copyright © 2003, Cisco Systems, Inc.


Configuring External Neighbors (Cont.)
Cisco.com
To temporarily disable a BGP neighbor:
router(config-router)#
neighbor ip-address shutdown
• Disables communication with a BGP neighbor
• Use scenarios:
– Debugging and troubleshooting
– Shutdown of the neighbor during extensive modification of routing policies to prevent inconsistent routing data
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-7

neighbor shutdown

To disable a neighbor, use the **neighbor shutdown** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **shutdown**

To re-enable the neighbor or peer group, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **shutdown**

Syntax Description

ip-address	IP address of neighbor
peer-group-name	Name of a BGP peer group

Practice

Q1) Which two parameters must you configure for a BGP neighbor? (Choose two.)

A) IP address

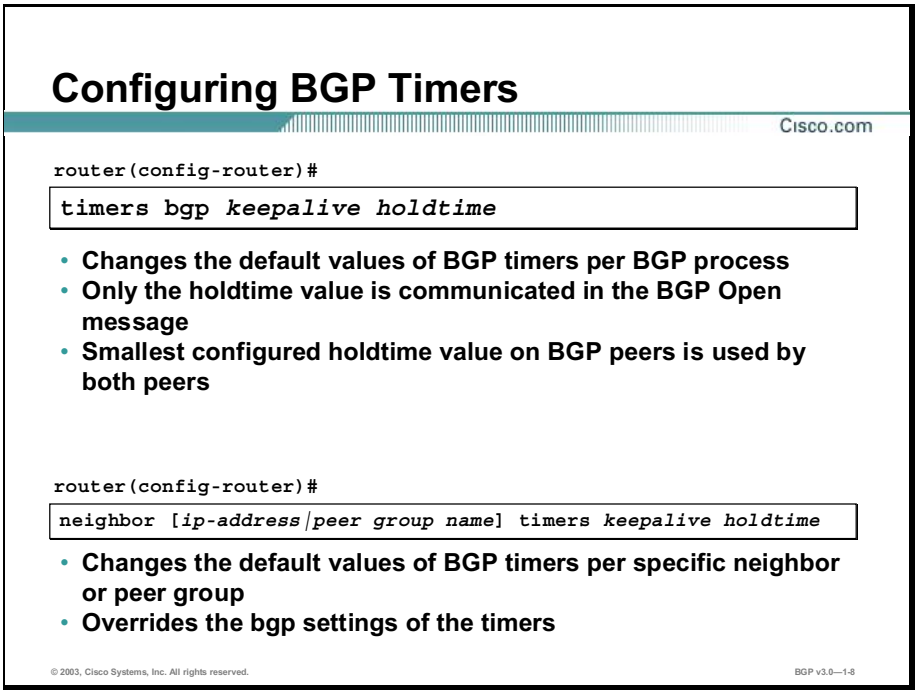
B) description

C) AS number

D) password

Configuring BGP Timers

This topic lists the commands required to modify the default keepalive and holdtime timers in BGP for the BGP process or for the TCP session between BGP neighbors.


Configuring BGP Timers
Cisco.com
router(config-router)#
timers bgp keepalive holdtime
• Changes the default values of BGP timers per BGP process
• Only the holdtime value is communicated in the BGP Open message
• Smallest configured holdtime value on BGP peers is used by both peers
router(config-router)#
neighbor [ip-address|peer group name] timers keepalive holdtime
• Changes the default values of BGP timers per specific neighbor or peer group
• Overrides the bgp settings of the timers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-8

Changing the BGP default holdtime and keepalive timers is usually not recommended. The defaults (keepalive: 60 seconds [sec], holdtime: 180 sec) should work fine in most situations. If for any reason a faster BGP response to a peer down event is needed (for example, in scenarios where multiple paths toward destinations are available), the neighbor timers on the router can be reduced. This reduction will result in a faster detection of a lost peer and faster switching to the alternate path in the BGP table, thus improving convergence.

A BGP router with an expired holdtime (no BGP traffic received inside the holdtime interval) will send a notification to its BGP peer, notifying it as to the reason for closing the session. The BGP router on which the holdtime has expired transitions the inactive peer into the “Idle” state. After a certain time interval, determined by auto-enable and connection timers, a BGP router again tries to reconnect to the previously disconnected BGP peer and will also accept connection attempts from that peer.

timers bgp

To adjust BGP network timers, use the **timers bgp** router configuration command.

timers bgp *keepalive holdtime*

To reset the BGP timing defaults, use the **no** form of this command.

no timers bgp

 Copyright © 2003, Cisco Systems, Inc.

Syntax Description

keepalive	Frequency (in seconds) with which the Cisco IOS software sends keepalive messages to its peer. The default is 60 seconds.
holdtime	Interval (in seconds) after not receiving a keepalive message that the software declares a peer dead. The default is 180 sec.

neighbor timers

To set the timers for a specific BGP peer or peer group, use the **neighbor timers** router configuration command. This command overrides the values set by the **timers bgp** command.

neighbor [*ip-address* | *peer-group-name*] **timers** *keepalive holdtime*

To clear the timers for a specific BGP peer or peer group, use the **no** form of this command.

no neighbor [*ip-address* | *peer-group-name*] **timers** *keepalive holdtime*

Syntax Description

ip-address	A BGP peer or peer group IP address
peer-group-name	Name of the BGP peer group
keepalive	Frequency (in seconds) with which the Cisco IOS software sends keepalive messages to its peer. The default is 60 sec.
holdtime	Interval (in seconds) after not receiving a keepalive message that the software declares a peer dead. The default is 180 sec.

Practice

Q1) What are the default values assigned to the keepalive and holdtime timers in BGP?

A) keepalive: 30 sec, holdtime: 90 sec

B) keepalive: 30 sec, holdtime: 120 sec

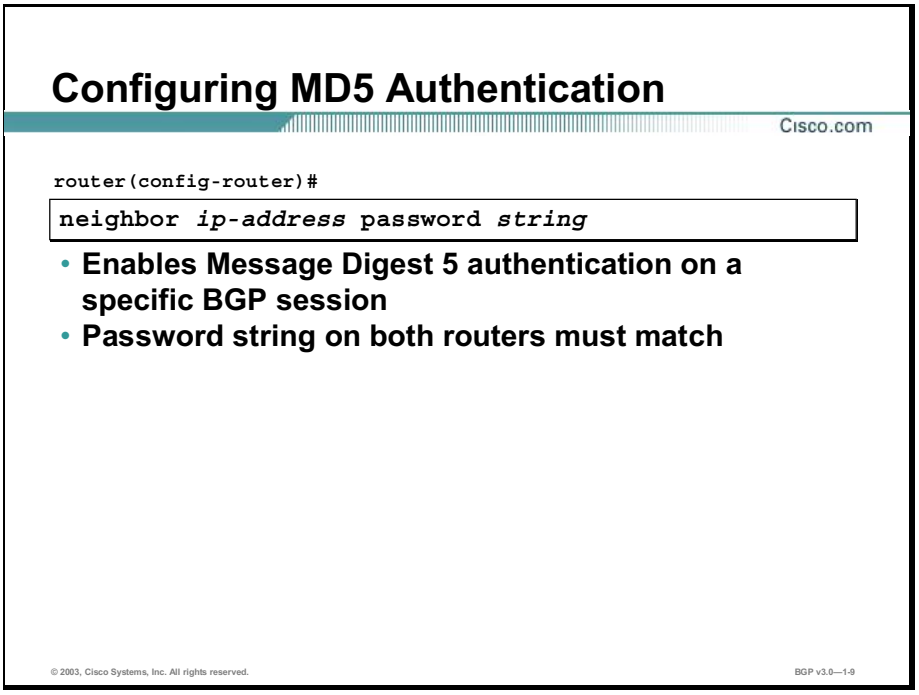
C) keepalive: 60 sec, holdtime: 180 sec

D) keepalive: 120 sec, holdtime: 270 sec

Copyright © 2003, Cisco Systems, Inc.

Configuring MD5 Authentication

This topic presents the command required to configure message digest authentication on a session between BGP neighbors.


Configuring MD5 Authentication
Cisco.com
router(config-router)#
neighbor ip-address password string
• Enables Message Digest 5 authentication on a specific BGP session
• Password string on both routers must match
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-9

neighbor password

To enable Message Digest 5 (MD5) authentication on a TCP connection between two BGP peers, use the **neighbor password** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **password** *string*

To disable this function, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **password**

Syntax Description

ip-address	IP address of neighbor
peer-group-name	Name of a BGP peer group
string	Case-sensitive password of up to 80 characters. The first character cannot be a number. The string can contain any alphanumeric characters, including spaces. You cannot specify a password in the format “number-space-anything.” The space after the number causes problems.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) In a BGP configuration using MD5 authentication, how must you configure passwords?

A) All routers in the AS must have the same password.

B) All neighbors on a single router must share the same password.

C) Neighbor routers must share the same password.

D) The hashing algorithm requires unique MD5 passwords.

Announcing Networks in BGP

This topic lists the Cisco IOS commands required to announce local networks to other BGP neighbors.

Before any local routing information can be injected by a router into its BGP table for advertising to other BGP routers, some basic configuration is required.

There are two different ways to do this configuration:

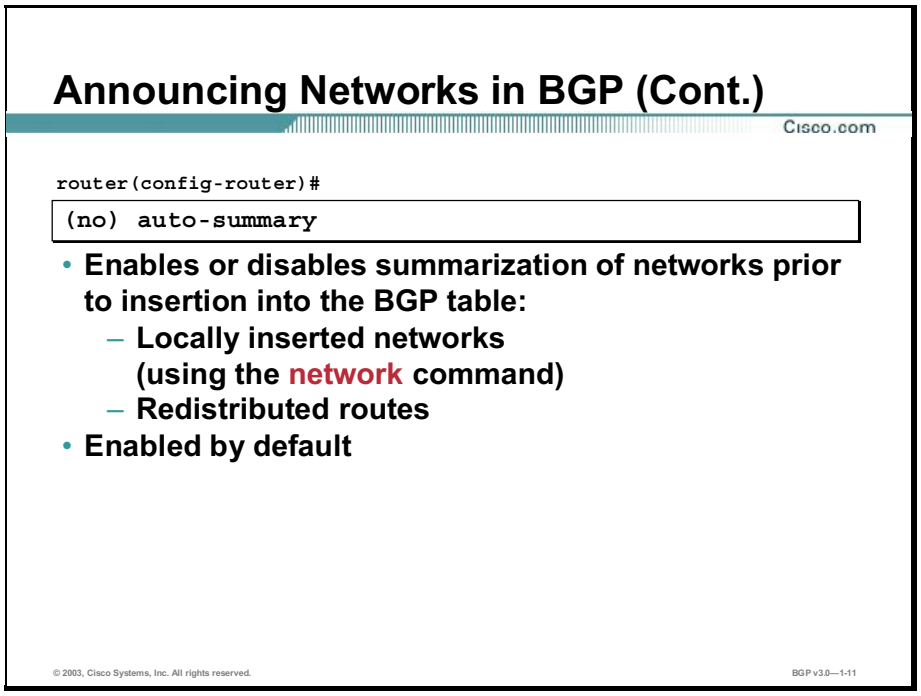
- List the network numbers that are candidates to be advertised. Do this with the **network** configuration command. If any of the listed networks are reachable by the local router, according to its forwarding table, then the network is injected as a route into the BGP table.

- Redistribute routing information learned by other routing protocols into the BGP table. You can use the Interior Gateway Protocol (IGP) used within the AS. Any route known by the local IGP can be injected into the BGP table using route redistribution between the IGP and BGP on the local router.

A router can also introduce new routing information into the BGP table by summarizing routes already there. This is called route aggregation and also requires configuration.

Any route introduced by the router into the BGP table will appear as a new route. The AS-path attribute for such a route will be empty, indicating a local route. The AS path changes later as the route passes AS boundaries.

 Copyright © 2003, Cisco Systems, Inc.


Announcing Networks in BGP (Cont.)
Cisco.com
router(config-router)#
(no) auto-summary
• Enables or disables summarization of networks prior to insertion into the BGP table:
– Locally inserted networks (using the network command)
– Redistributed routes
• Enabled by default
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-11

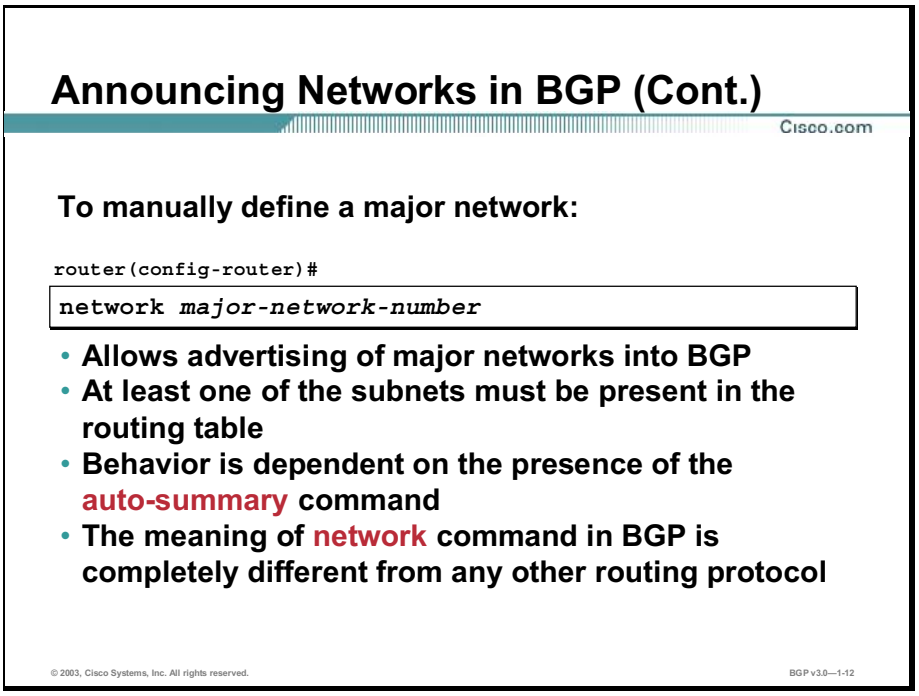
When the router is configured to locally announce routes into BGP, the behavior of the **network** command varies depending on whether automatic summarization is enabled or disabled. When automatic summarization is enabled, it summarizes locally originated BGP networks to their classful boundaries. By default, automatic summarization is enabled for BGP.

When a subnet exists in the routing table and the following three conditions are satisfied, then any subnet (component route) of that classful network in the local routing table will prompt BGP to install the classful network into the BGP table:

- A classful **network** statement for a network that exists in the routing table.
- A classful mask has been configured on that **network** statement.
- Automatic summarization is enabled.

When automatic summarization is disabled, the routes introduced locally into the BGP table are not summarized to their classful boundaries.

The BGP **auto-summary** command is also responsible for the behavior of the redistribution procedure in BGP. When enabled, all redistributed subnets will be summarized to their classful boundaries in the BGP table. When disabled, all redistributed subnets will be present in their original form in the BGP table.

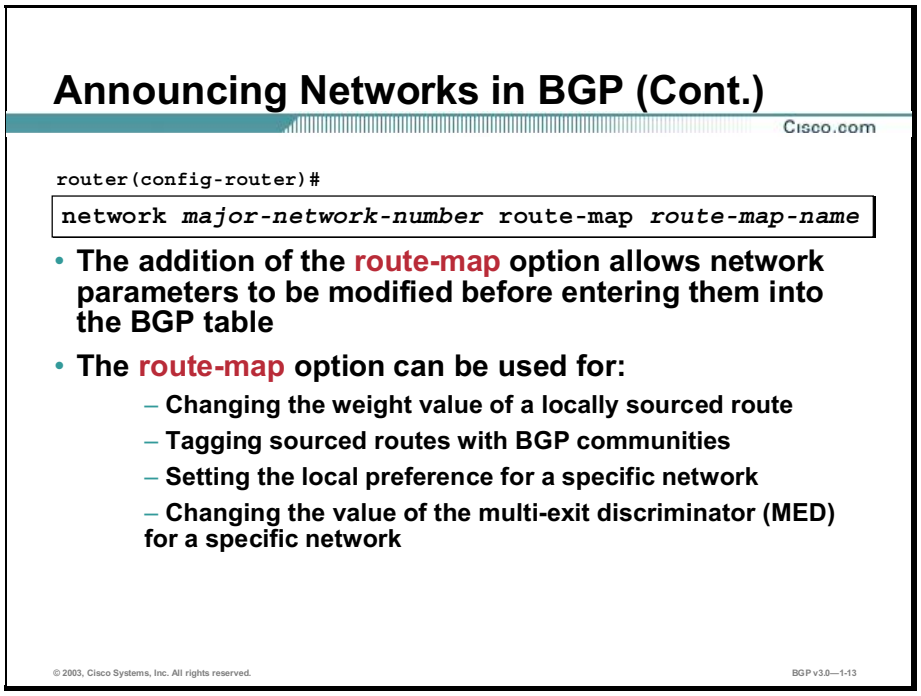

Announcing Networks in BGP (Cont.)
Cisco.com
To manually define a major network:
router(config-router)#
network major-network-number
• Allows advertising of major networks into BGP
• At least one of the subnets must be present in the routing table
• Behavior is dependent on the presence of the auto-summary command
• The meaning of network command in BGP is completely different from any other routing protocol
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-12

To specify the networks to be advertised by the BGP routing process, use the **network** router configuration command. To remove an entry, use the **no** form of this command.

Note	The meaning of the **network** command in BGP is radically different from the way that the command is used in other routing protocols. In all other routing protocols, the **network** command indicates interfaces over which the routing protocol will be run. In BGP, it indicates only which routes should be injected into the BGP table on the local router. Also, BGP never runs over individual interfaces; it is run over TCP sessions with manually configured neighbors.

The **network** command with no **mask** option uses the classful approach to insert a major network into the BGP table. At least one subnet of the specified major network needs to be present in the IP routing table to allow BGP to start announcing the major network as a BGP route. If automatic summarization is disabled, an exact match is required.

 Copyright © 2003, Cisco Systems, Inc.


Announcing Networks in BGP (Cont.)
Cisco.com
router(config-router)#
network major-network-number route-map route-map-name
• The addition of the route-map option allows network parameters to be modified before entering them into the BGP table
• The route-map option can be used for:
– Changing the weight value of a locally sourced route
– Tagging sourced routes with BGP communities
– Setting the local preference for a specific network
– Changing the value of the multi-exit discriminator (MED) for a specific network
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-13

When the router is configured to insert routes in the BGP table, the default attributes of locally sourced routes can be modified with the inclusion of the **route-map** option into the basic **network** command.

The attached route-map can change the following attributes of locally sourced networks with the **network** command:

- Weight (default value=32768): The weight attribute is a special Cisco attribute that is used in the path selection process when there is more than one route to the same destination. Because weight is considered before local preference in the BGP route selection, locally sourced routes are always preferred, unless weight value is modified.
- Community (default value=nonexistent): Used for tagging routes at their source.
- Local preference (default value=100): Used for AS-wide BGP best-path selection.
- Multi-exit discriminator (MED; default value=0): Used for return path selection in topologies, where multiple exit points to the same neighbor AS exist.

Example

If a subnet existing in the routing table is 75.75.75.0 mask 255.255.255.0, and network 75.0.0.0 is configured under the **router bgp** command (assuming automatic summarization is enabled), BGP will introduce the classful network 75.0.0.0 mask 255.0.0.0 in the BGP table. If the following three conditions are not all met, then BGP will not install any entry in the BGP table unless there is an exact match in the IP routing table.

- A classful **network** statement for a network that exists in the routing table.
- A classful mask has been configured on that **network** statement.
- Automatic summarization is enabled.

Practice

Q1) How is the **network** command in BGP different from other routing protocols?

A) The BGP **network** command signifies what router interfaces should run BGP.

B) The **network** command in BGP contains a mask.

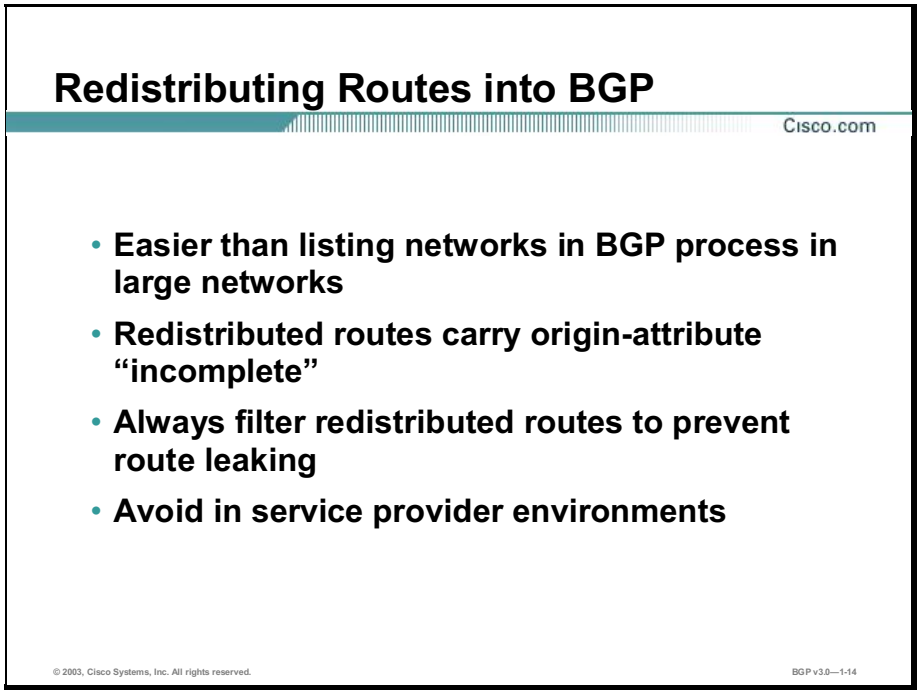
C) The **network** command indicates what routes should be injected into the BGP table.

D) The **network** command is the same for all IP routing protocols.

 Copyright © 2003, Cisco Systems, Inc.

Redistributing Routes into BGP

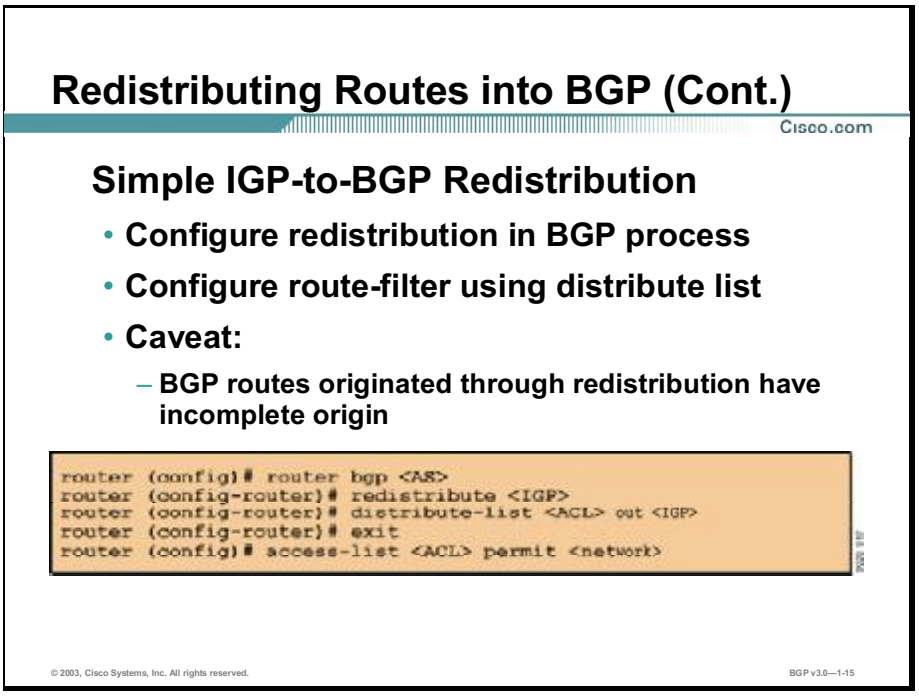
This topic describes route redistribution in BGP and identifies the Cisco IOS commands required to configure BGP route redistribution.


Redistributing Routes into BGP
Cisco.com
• Easier than listing networks in BGP process in large networks
• Redistributed routes carry origin-attribute "incomplete"
• Always filter redistributed routes to prevent route leaking
• Avoid in service provider environments
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-14

There are two alternatives for injecting local routes into the BGP table: list them using the **network** command or redistribute them. Listing the routes gives you total control over networks that could possibly be advertised by BGP. This is a very desirable option for multihomed customers or Internet service providers (ISPs). On the other hand, this approach requires a lot of configuration commands that could be hard to maintain.

If there are a lot of networks to be advertised, and BGP is used primarily to achieve scalability, not routing security (for example, in enterprise networks), it could be easier to let the local IGP find the routes and then redistribute them into BGP. However, this approach introduces the risk that the IGP may find some networks that were not supposed to be advertised. Private network numbers, such as network 10.0.0.0/8, are often used within an AS for various reasons but must never be advertised out to the Internet. Careful filtering must be done to prevent unintentional advertising.

When the router injects a route listed with a **network** command into its BGP table, the origin code is set to "IGP." If the route is injected into the BGP table through redistribution, the origin code is set to "unknown/incomplete."


Redistributing Routes into BGP (Cont.)
Cisco.com
Simple IGP-to-BGP Redistribution
• Configure redistribution in BGP process
• Configure route-filter using distribute list
• Caveat:
– BGP routes originated through redistribution have incomplete origin
router (config)# router bgp <AS>
router (config-router)# redistribute <IGP>
router (config-router)# distribute-list <ACL> out <IGP>
router (config-router)# exit
router (config)# access-list <ACL> permit <network>
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-15

Routes redistributed into BGP will carry the origin attribute “Incomplete”. In most cases this situation does not jeopardize BGP functionality. It could pose a problem if the route selection process has to decide on the best route toward a particular destination based on the MED attribute. In the case of receiving two routes, one with the “IGP” origin (inserted with the **network** command), and another one with “Incomplete” origin, the first route would always be selected, no matter what value the MED attribute is set to (according to the BGP route selection rules).

redistribute (IP)

To redistribute routes from one routing process into another routing process, use the **redistribute** router configuration command.

redistribute *protocol* [*process-id*] {**level-1** | **level-1-2** | **level-2**} [**metric** *metric-value*] [**metric-type** *type-value*] [**match** {**internal** | **external 1** | **external 2**}] [**tag** *tag-value*] [**route-map** *map-tag*] [**weight** *weight*] [**subnets**]

To disable redistribution, use the **no** form of this command.

no redistribute *protocol* [*process-id*] {**level-1** | **level-1-2** | **level-2**} [**metric** *metric-value*] [**metric-type** *type-value*] [**match** {**internal** | **external 1** | **external 2**}] [**tag** *tag-value*] [**route-map** *map-tag*] [**weight** *weight*] [**subnets**]

 Copyright © 2003, Cisco Systems, Inc.

Syntax Description

protocol	Source protocol from which routes are being redistributed. It can be one of the following keywords: **bgp**, **egp**, **igrp**, **isis**, **ospf**, **static** [*ip*], **connected**, and **rip**. The keyword **connected** refers to routes that are established automatically by virtue of having enabled IP on an interface.
process-id	(Optional) For **bgp**, **egp**, or **igrp**, this is an AS number, which is a 16-bit decimal number. For **isis**, this is an optional tag that defines a meaningful name for a routing process. You can specify only one Intermediate System-to-Intermediate System (IS-IS) process per router. Creating a name for a routing process means that you use names when configuring routing. For **ospf**, this is an appropriate Open Shortest Path First (OSPF) process ID from which routes are to be redistributed. This ID identifies the routing process. This value takes the form of a nonzero decimal number. For **rip**, no process ID value is needed.
level-1	For IS-IS, Level 1 routes are redistributed into other IP routing protocols independently.
level-1-2	For IS-IS, both Level 1 and Level 2 routes are redistributed into other IP routing protocols.
level-2	For IS-IS, Level 2 routes are redistributed into other IP routing protocols independently.
metric *metric-value*	(Optional) Metric used for the redistributed route. If a value is not specified for this option, and no value is specified using the **default-metric** command, the default metric value is 0. In the case of BGP, the metric will be the MED value.
match {**internal** | **external 1** | **external 2**}	(Optional) For OPSF, the criteria by which OSPF routes are redistributed into other routing processes. It can be one of the following: internal—Routes that are internal to a specific AS. external 1—Routes that are external to the AS but are imported into OSPF as a type 1 external route. external 2—Routes that are external to the AS but are imported into OSPF as a type 2 external route.

route-map	(Optional) The route-map should be interrogated to filter the importation of routes from this source routing protocol to the current routing protocol. If not specified, all routes are redistributed. If this keyword is specified, but no route-map tags are listed, no routes will be imported.
map-tag	(Optional) Identifier of a configured route-map.
weight *weight*	(Optional) Network weight when redistributing into BGP. An integer from 0 to 65535.
subnets	Indicates that not only networks with a natural mask should be redistributed but also subnets.

distribute-list out (IP)

To suppress networks from being advertised in updates, use the **distribute-list out** router configuration command with the *routing-process* specified.

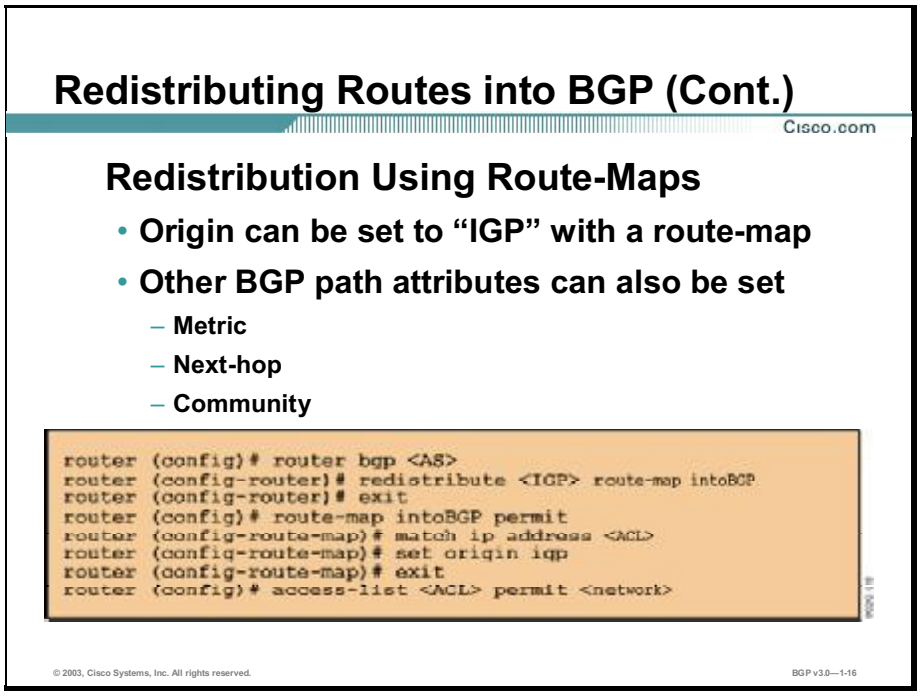
distribute-list {*access-list-number* | *access-list-name*} **out** [*interface-name* | *routing-process* | *autonomous-system-number*]

To cancel this function, use the **no** form of this command.

no distribute-list {*access-list-number* | *access-list-name*} **out** [*interface-name* | *routing-process* | *autonomous-system-number*]

The access-list referred to by the **distribute-list** command permits those routes that should be redistributed.

 Copyright © 2003, Cisco Systems, Inc.


Redistributing Routes into BGP (Cont.)
Cisco.com
Redistribution Using Route-Maps
• Origin can be set to "IGP" with a route-map
• Other BGP path attributes can also be set
– Metric
– Next-hop
– Community
router (config)# router bgp <AS>
router (config-router)# redistribute <IGP> route-map intoBGP
router (config-router)# exit
router (config)# route-map intoBGP permit
router (config-route-map)# match ip address <ACL>
router (config-route-map)# set origin igp
router (config-route-map)# exit
router (config)# access-list <ACL> permit <network>
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-16

Route-maps can be configured on the router to filter updates and modify various attributes. A configured route-map can be applied to routes being redistributed from the IGP.

Only the routes permitted by the route-map will be redistributed. Using the **set** command in the route-map, specific path attributes attached to the redistributed routes can be modified. Thus, only selected routes will be advertised, and they will have the desired attribute values.

The route-map must be given a name. This name is a case-sensitive string, which is used when referring to it. Any string could be used but a meaningful name is suggested.

Use the **route-map** global configuration command, and the **match** and **set** route-map configuration commands, to define the conditions for redistributing routes. Each **route-map** command has a list of **match** and **set** commands associated with it. The **match** commands specify the match criteria—the conditions under which redistribution is allowed for the current **route-map** command. The **set** commands specify the set actions—the particular redistribution actions to perform if the criteria enforced by the **match** commands are met.

When you are passing routes through a route-map, it can have several parts. Any route that does not match at least one match clause relating to a **route-map** command will be ignored; that is, the route will not be advertised. If you only want to modify some data, you must configure a second route-map section with an explicit match specified.

Practice

Q1) What is a potential problem caused by redistributing all routes from the IGP to BGP?

A) Private addresses (RFC 1918) or other networks that the local AS does not wish to advertise will be propagated into BGP.

B) Summary prefixes will be announced, creating routing “black holes.”

C) The BGP neighbor will be unable to apply incoming route filters.

D) The BGP routing protocol process will become unstable due to the faster convergence requirements of the IGP routes.

Q2) What is a key advantage of using route-maps when redistributing IGP routes into BGP?

A) The weight attribute can be set.

B) Redistribution of routes can be matched to an access-list.

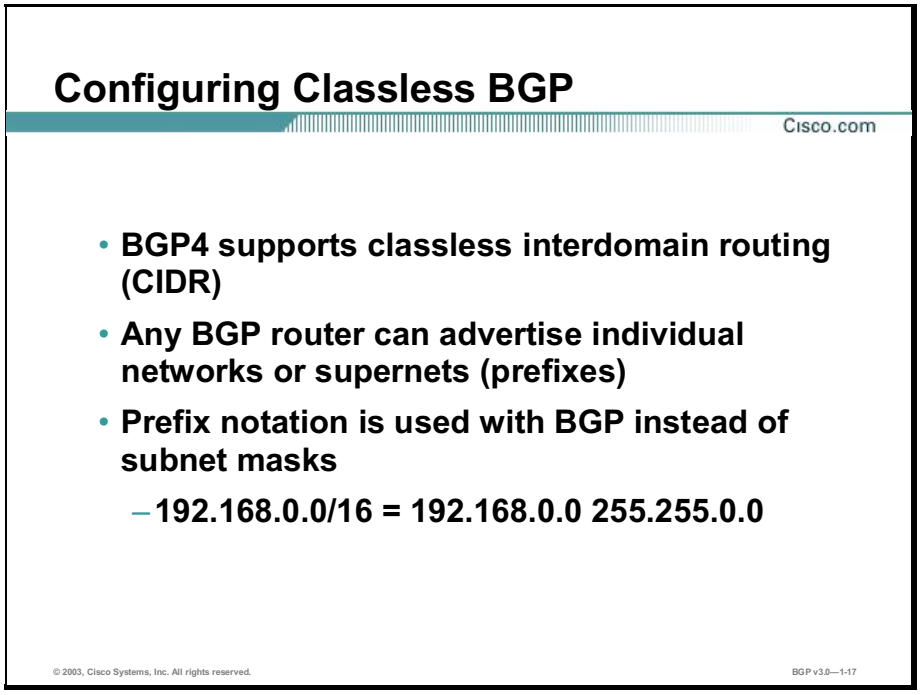
C) The direction of the redistribution can be controlled.

D) The origin attribute can be set to IGP.

 Copyright © 2003, Cisco Systems, Inc.

Configuring Classless BGP

This topic describes the classless behavior of BGP and the commands required to advertise a classless BGP supernet prefix.

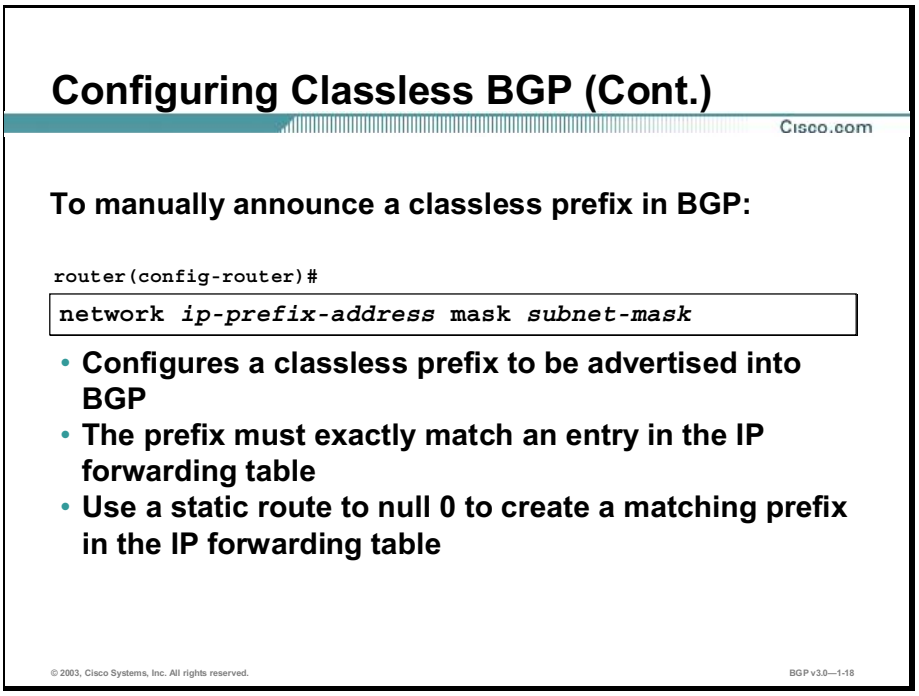

Configuring Classless BGP
Cisco.com
• BGP4 supports classless interdomain routing (CIDR)
• Any BGP router can advertise individual networks or supernets (prefixes)
• Prefix notation is used with BGP instead of subnet masks
– 192.168.0.0/16 = 192.168.0.0 255.255.0.0
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-17

BGP version 4 is a classless protocol, meaning that its routing updates include IP address and the subnet mask. The combination of the IP address and the subnet mask is called an IP prefix. An IP prefix can be a subnet, a major network, or a supernet.

BGP uses prefix notation (address/number of bits) to display IP prefixes. The number following the slash, "/", in the 192.168.0.0/16 notation in the figure is referring to the number of bits in the subnet mask being set to 1. The subnet mask 255.255.0.0 starts with 16 consecutive bits set to 1, and the rest of the bits set to zero.

As another example, the subnet 172.16.1.0 with mask 255.255.255.0 can be written using the prefix notation as 172.16.1.0/24.

When classless prefix notation is used, an old class A network, for example, 10.0.0.0, with the natural mask, is written as 10.0.0.0/8. A class B network, 172.17.0.0 with natural mask, is written as 172.17.0.0/16, and a class C network, 192.168.1.0 with natural mask, is written as 192.168.1.0/24.


Configuring Classless BGP (Cont.)
Cisco.com
To manually announce a classless prefix in BGP:
router(config-router)#
network ip-prefix-address mask subnet-mask
• Configures a classless prefix to be advertised into BGP
• The prefix must exactly match an entry in the IP forwarding table
• Use a static route to null 0 to create a matching prefix in the IP forwarding table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-18

To advertise classless networks into BGP (a subnet or a supernet), you can use the **network** command with the **mask** keyword and the subnet mask specified. When an exact match is not found in the IP routing table (for example, when creating a summary or when advertising only a part of your address space), a matching prefix has to be manually configured on the router in the form of a static route pointing to the null 0 interface; otherwise the advertisement will not succeed.

network (BGP)

To specify the networks to be advertised by the BGP routing process, use the **network** router configuration command.

network *network-number* [**mask** *network-mask*]

To remove an entry, use the **no** form of this command.

no network *network-number* [**mask** *network-mask*]

Syntax Description

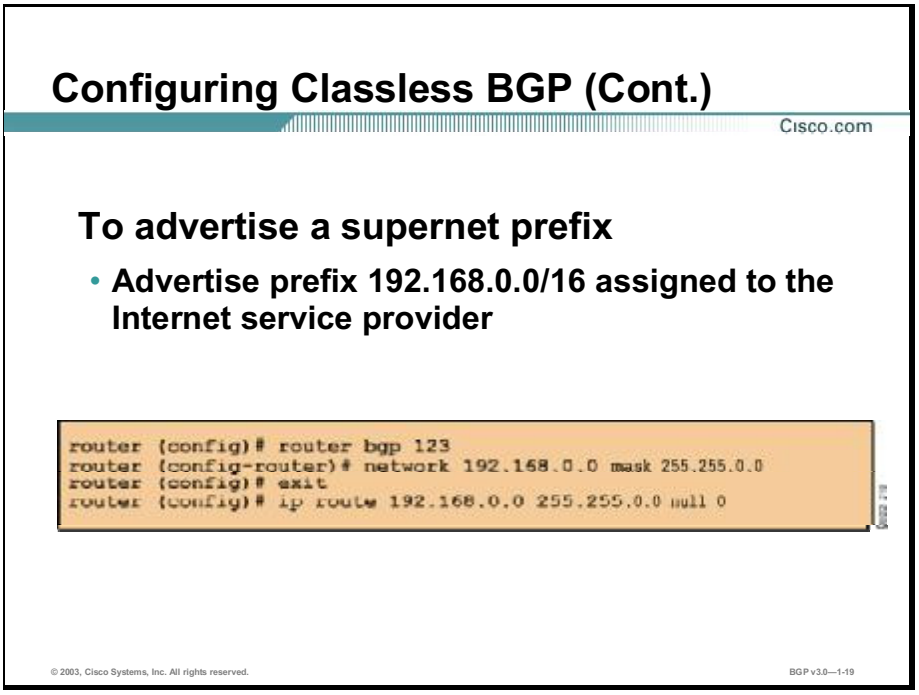
network-number	Network that BGP will advertise
mask	(Optional)
network-mask	(Optional) Network mask address

 Copyright © 2003, Cisco Systems, Inc.

If the keyword **mask** and the subnet mask are omitted, the network is assumed to have its natural mask according to the network class. In this case, the route will still be injected into the BGP table on the router if there is any subnet of the major network that is reachable according to the forwarding table.

If the network mask is specified, the behavior changes slightly and it is required that an exact match of network number and subnet mask appear in the forwarding table before the route is injected into the BGP table.

Example


Configuring Classless BGP (Cont.)
Cisco.com
To advertise a supernet prefix
• Advertise prefix 192.168.0.0/16 assigned to the Internet service provider
router (config)# router bgp 123
router (config-router)# network 192.168.0.0 mask 255.255.0.0
router (config)# exit
router (config)# ip route 192.168.0.0 255.255.0.0 null 0
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-19

In this example, the IP address space 192.168.0.0/16 is assigned to a service provider, and the service provider would like that address space to be constantly advertised by BGP. The **network** command with the **mask** option tells BGP that 192.168.0.0/16 is a candidate for being advertised. The **mask** keyword and the mask 255.255.0.0 are required because the mask is not the natural one.

However, before the candidate route is actually advertised, the router checks the forwarding table for an exact match (both network number and mask). It will always be found because there is a static route for it. This static route points to the null interface, which is always available.

The conclusion is that 192.168.0.0/16 will always be advertised by this router. All other BGP routers will use this information and forward any IP packets with the destination IP address in the interval 192.168.0.0 to 192.168.255.255 (inclusive) in the direction of this router.

When those packets arrive, the router, in this example, must have more explicit routes to the different parts of the 192.168.0.0/16 address range. This need could be answered by the IGP, which is not shown in the configuration example.

If, however, an IP packet arrives with a destination address to which this router does not have a more explicit route, the static route will route the packet to the null interface where it is dropped. This routing is a safety precaution that will prevent a routing loop, which might occur when route summaries are used in combination with default routing. If a packet arrives from the Internet to a subnet of 192.168.0.0/16, which is currently not reachable, the packet might otherwise have followed the default route toward the Internet because there was no more explicit route. Of course, the packet would immediately be routed back again, and a routing loop would have occurred.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) In BGP prefix notation, what does the number following the slash (/) indicate?

A) the number of bits in the subnet mask being set to 1

B) the number of subnet bits following the natural mask

C) the number of bits in the host portion of an IP address

D) the mask to use in the octet of subnetting

Q2) If you use the **network mask** option with the BGP **network** command, what must be present for the aggregate to appear in the BGP routing table?

A) a loopback interface on the router in the same network as the summary

B) an exact match of the network number and mask in the IP routing table

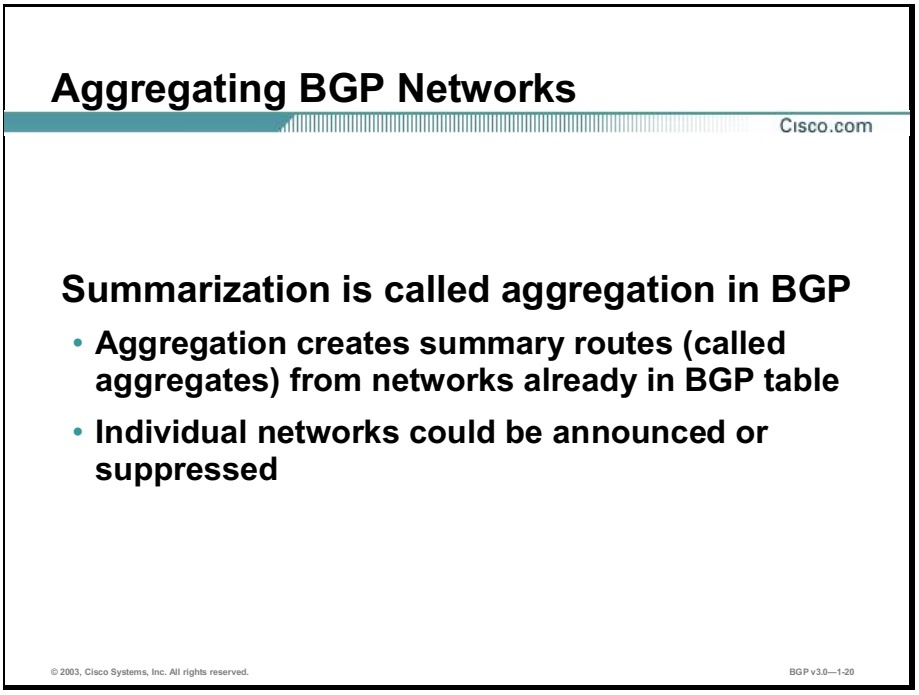
C) specification of the **summary-only** option

D) a default route to null 0 on the router

Copyright © 2003, Cisco Systems, Inc.

Aggregating BGP Networks

This topic describes route summarization in BGP. It also lists the configuration commands required to configure summary routes in BGP.


Aggregating BGP Networks
Cisco.com
Summarization is called aggregation in BGP
• Aggregation creates summary routes (called aggregates) from networks already in BGP table
• Individual networks could be announced or suppressed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-20

When the BGP table is already populated with routes that should be summarized, you must configure a router to do so. The summarization of BGP routes is called aggregation.

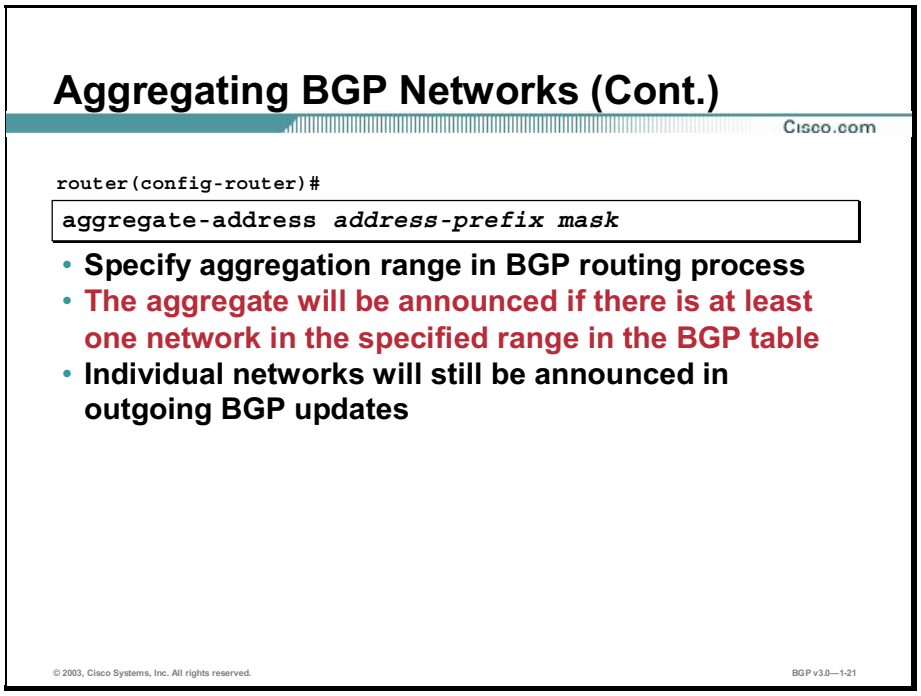
Use aggregation when a group of more specific routes has been injected into the BGP table at one stage but can be summarized at a later stage. The routes to be summarized could be IGP routes redistributed into BGP. Before BGP advertises these routes to the rest of the network, an aggregation of the subnets into a larger announcement would be appropriate.

In some networks, more specific routes are injected into the BGP table by some routers and aggregation is done in another router or even in another AS. This is called proxy aggregation.

When a router is configured to do aggregation, you must configure the route summary. If any route already in the BGP table is within the range indicated by the summary, then the summary route is also injected into the BGP table on the route and advertised to other routers. This action creates more information in the BGP table. To get any benefits from the aggregation, you must suppress the more specific routes, which are covered by the route summary. This suppression is an option to the **aggregate** configuration command.

When you suppress the more specific routes through configuration, they are still present in the BGP table of the router doing the aggregation. However, because the routes are marked as suppressed, they are never advertised to any other router.

 Copyright © 2003, Cisco Systems, Inc.


Aggregating BGP Networks (Cont.)
Cisco.com
router(config-router)#
aggregate-address address-prefix mask
• Specify aggregation range in BGP routing process
• The aggregate will be announced if there is at least one network in the specified range in the BGP table
• Individual networks will still be announced in outgoing BGP updates
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-21

In this configuration command syntax, where the keyword **summary-only** is not used, both the route summary and the more specific routes will be advertised. This approach is generally not desirable. Therefore, the suppressing of individual routes, described next, is used in most cases.

aggregate-address

To create an aggregate entry in a BGP routing table, use the **aggregate-address** router configuration command.

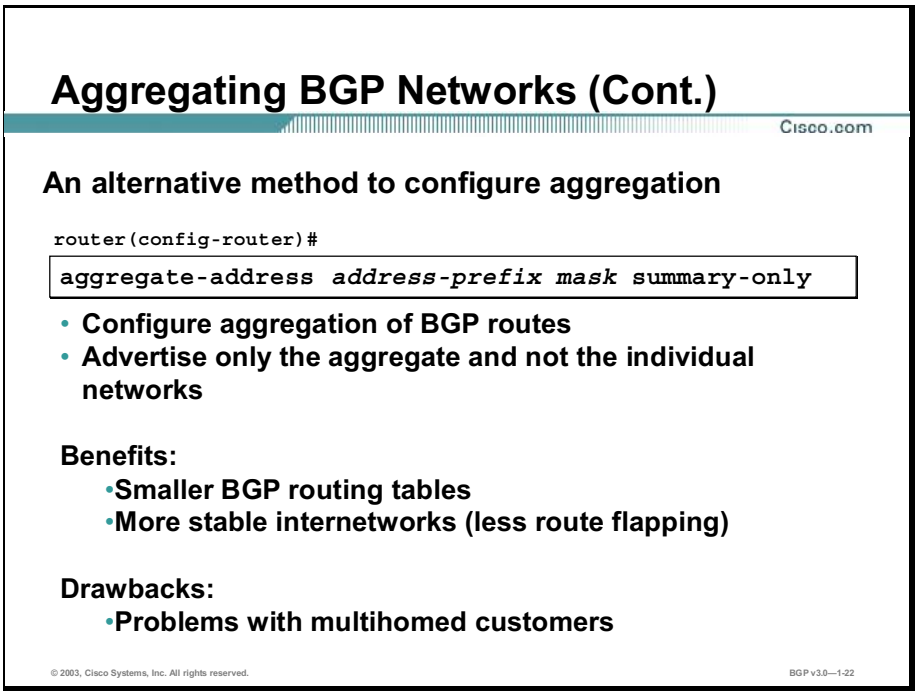
aggregate-address *address mask* [**as-set**] [**summary-only**] [**suppress-map** *map-name*][**advertise-map** *map-name*] [**attribute-map** *map-name*]

To disable this function, use the **no** form of this command.

no aggregate-address *address mask* [**as-set**] [**summary-only**] [**suppress-map** *map-name*][**advertise-map** *map-name*] [**attribute-map** *map-name*]

Syntax Description

address	Aggregate address
mask	Aggregate mask
summary-only	(Optional) Suppresses more specific routes


Aggregating BGP Networks (Cont.)
Cisco.com
An alternative method to configure aggregation
router(config-router)#
aggregate-address address-prefix mask summary-only
• Configure aggregation of BGP routes
• Advertise only the aggregate and not the individual networks
Benefits:
•Smaller BGP routing tables
•More stable internetworks (less route flapping)
Drawbacks:
•Problems with multihomed customers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-22

When the **summary-only** option is used, only the route summary will be advertised, not the more specific routes.

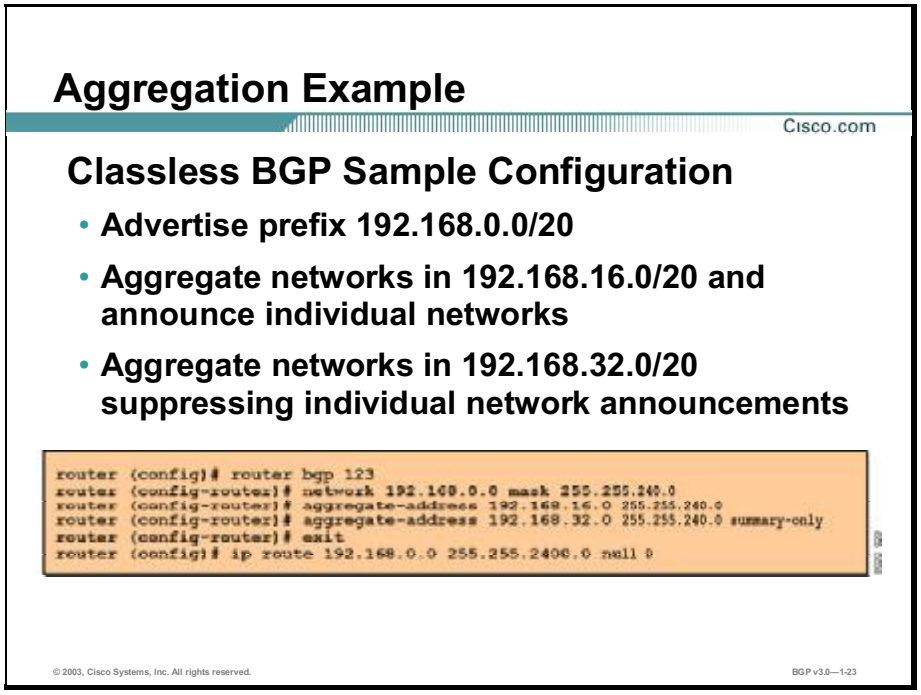
One of the benefits of this approach is that the rest of the routers will receive only one route instead of many more specific routes. It eases the burden on the other routers by reducing the amount of memory required to hold the BGP table.

Another benefit is that route flapping is reduced. The router doing the aggregation will keep on advertising the aggregate as long as there is at least one specific route within the range still available. If one of the more specific routes is lost, but at least one remains, the aggregate itself will not be lost. The flap of the more specific route is not visible to the rest of the network. This approach reduces the amount of updates necessary and the CPU power required to process them.

However, all route summarization in any routing protocol causes a loss of granularity. Suboptimal routing could be introduced when redundant paths are available to reach a group of networks advertised by a single route summary. Some of the networks could be more reachable via one of the paths, while others may be more reachable another way. From outside the immediate network, multiple paths may not be visible because only summary routes are advertised. Therefore, there is a risk that the least optimum path will be chosen.

 Copyright © 2003, Cisco Systems, Inc.

Example


Aggregation Example
Cisco.com
Classless BGP Sample Configuration
• Advertise prefix 192.168.0.0/20
• Aggregate networks in 192.168.16.0/20 and announce individual networks
• Aggregate networks in 192.168.32.0/20 suppressing individual network announcements
router (config)# router bgp 123
router (config-router)# network 192.168.0.0 mask 255.255.240.0
router (config-router)# aggregate-address 192.168.16.0 255.255.240.0
router (config-router)# aggregate-address 192.168.32.0 255.255.240.0 summary-only
router (config-router)# exit
router (config)# ip route 192.168.0.0 255.255.240.0 null 0
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-23

The configuration example here shows three different ways of advertising a route summary. This example will be used to describe basic BGP monitoring commands.

- The prefix 192.168.0.0/20 is always advertised. It is injected into the BGP table as a summary. The **network** statement makes it a candidate for being advertised. Because the mask is specified, an exact match in the routing table is a required condition before the route is injected into the BGP table. The matching route is inserted in the IP routing table by the static **ip route** statement to the null 0 interface.

- The prefix 192.168.16.0/20 is conditionally advertised. It is injected into the BGP table whenever there is a more specific route within the route summary range already in the BGP table. However, the more specific route is still advertised.

- The prefix 192.168.32.0/20 is also conditionally advertised. It is injected into the BGP table whenever there is a more specific route within the route summary range already in the BGP table. However, any more specific routes are suppressed and not advertised to any neighbors.

Aggregation Example (Cont.)

Cisco.com

Viewing the BGP table

```
as123#show ip bgp
BGP table version is 16, local router ID is 1.2.3.4
Status codes: s suppressed, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop        Metric  LocPrf  Weight  Path
*> 1.0.0.0          0.0.0.0              0           32768  i
*  21.0.0.0         3.4.5.6                             0  37 21 i
*>                  2.3.4.5              0              0  21 i
*> 37.0.0.0         3.4.5.6              0              0  37 i
*                   2.3.4.5                             0  21 37 i
*> 192.168.0.0/20   0.0.0.0              0           32768  i
*> 192.168.16.0/20  0.0.0.0              0           32768  i
*> 192.168.16.0     0.0.0.0              0           32768  ?
*> 192.168.17.0     0.0.0.0              0           32768  ?
*> 192.168.32.0/20  0.0.0.0              0           32768  i
s> 192.168.32.0     0.0.0.0              0           32768  ?
s> 192.168.33.0     0.0.0.0              0           32768  ?
```

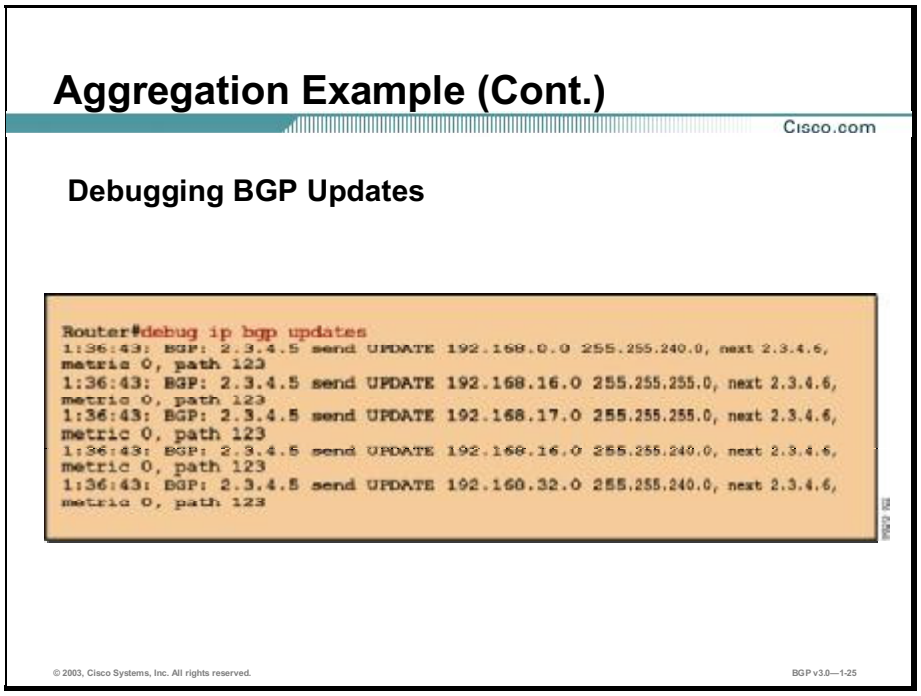
© 2003, Cisco Systems, Inc. All rights reserved.

The **show ip bgp** command prints the BGP table. As shown above, all three prefixes are injected:

- The prefix 192.168.0.0/20 is always injected.
- The prefix 192.168.16.0/20 is injected because there is at least one more specific route within the summary range. In this case, both 192.168.16.0/24 and 192.168.17.0/24 are within the range. Nothing is changed with the more specific routes, so they are still advertised.
- The prefix 192.168.32.0/20 is injected because there is at least one more specific route within the summary range. In this case, both 192.168.32.0/24 and 192.168.33.0/24 are within the range. The more specific routes are marked as suppressed using the lowercase letter “s.” The “s” means that they are still present and available in the BGP table of the router, but they are not advertised on any BGP session.

Note Because the prefixes 192.168.16.0/24, 192.168.17.0/24, 192.168.32.0/24, and 192.168.33.0/24 all have natural masks as applied to class C networks, the prefix length is not displayed on the **show ip bgp** printout. The network mask is, however, stored in the BGP table and sent on any BGP update.

 Copyright © 2003, Cisco Systems, Inc.


Aggregation Example (Cont.)
Cisco.com
Debugging BGP Updates
Router#debug ip bgp updates
1:36:43: BGP: 2.3.4.5 send UPDATE 192.168.0.0 255.255.240.0, next 2.3.4.6, metric 0, path 123
1:36:43: BGP: 2.3.4.5 send UPDATE 192.168.16.0 255.255.255.0, next 2.3.4.6, metric 0, path 123
1:36:43: BGP: 2.3.4.5 send UPDATE 192.168.17.0 255.255.255.0, next 2.3.4.6, metric 0, path 123
1:36:43: BGP: 2.3.4.5 send UPDATE 192.168.16.0 255.255.240.0, next 2.3.4.6, metric 0, path 123
1:36:43: BGP: 2.3.4.5 send UPDATE 192.160.32.0 255.255.240.0, next 2.3.4.6, metric 0, path 123
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-25

The debug output shows the BGP updates that have been sent to a neighbor. All three route summary prefixes, 192.168.0.0/20, 192.168.16.0/20 and 192.168.32.0/20, are included in the updates. Also, the nonsuppressed more explicit routes, 192.168.16.0/24 and 192.168.17.0/24, are included in the update. However, the suppressed more explicit routes, 192.168.32.0/24 and 192.168.33.0/24, are never sent as updates on the BGP session.

Practice

Q1) What is the purpose of the **summary-only** option when you are configuring BGP route aggregation?

A) to explicitly include summary routes in route aggregates

B) to summarize BGP routes to their classful boundaries

C) to suppress the advertising of routes more specific than the aggregate address

D) to prevent summary routes from being advertised within an AS

Q2) What are two potential drawbacks to using route summarization? (Choose two.)

A) causes loss of granular routing information

B) reduces the size of the routing table

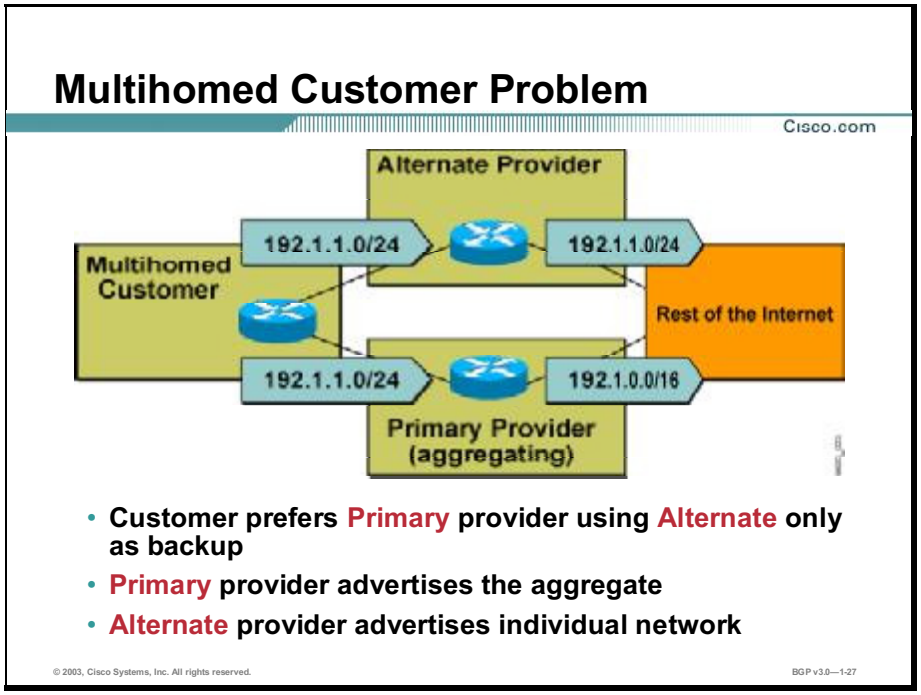
C) prohibits redistribution of summarized routes

D) can cause suboptimal routing to occur

 Copyright © 2003, Cisco Systems, Inc.

Multihomed Customer Problem

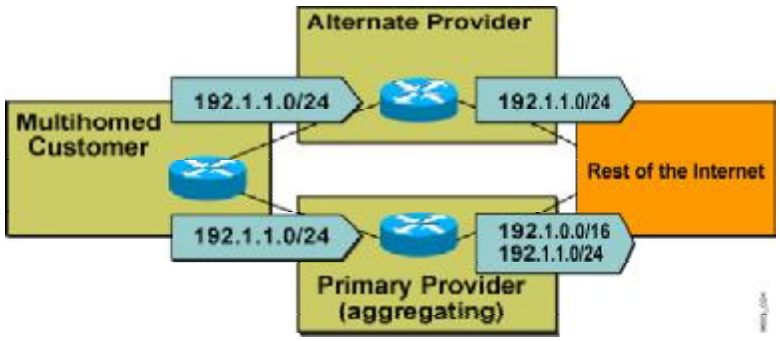
This topic describes when route aggregation in BGP is not appropriate.



In this example, the primary provider is doing aggregation of 192.1.0.0/16 before sending it to the rest of the network. This situation means that the primary provider is also doing proxy aggregation for the route 192.1.1.0/24 advertised by the multihomed customer. The rest of the Internet will not see the route 192.1.1.0/24 via the primary provider.

But the multihomed customer also advertises 192.1.1.0/24 to the alternate provider. In this case, the provider does not do any aggregation of any routes starting with 192.1 (and should not do so). This situation means that the alternate provider will propagate 192.1.1.0/24 to the rest of the Internet.

Multihomed Customer Problem (Cont.)



- **Remote autonomous systems prefer longest-match prefix; traffic toward the customer flows through Alternate provider**
- **Solution: Don't use aggregation**

© 2003, Cisco Systems, Inc. All rights reserved.

The rest of the Internet now sees overlapping routes. It sees 192.1.1.0/24 reachable via the alternate provider and 192.1.0.0/16 reachable via the primary provider. These two routes are treated as different routes. They are not compared with each other in a route selection process because they indicate different destinations. Because the router views them as different destinations, both routes will be injected into the forwarding table.

If a packet arrives with a destination address in the 192.1.1.0/24 network, the rest of the Internet will follow the "longest matching prefix" rule and forward the packet to the alternate provider.

To avoid this, the primary provider must turn off aggregation. If the primary provider does so, the rest of the Internet will see 192.1.1.0/24 both ways. And, because exactly the same route (network and mask) is reachable two ways, route-selection processing starts. Depending on the attribute values, the rest of the Internet could be advised to use the primary provider instead of the alternate one.

However, turning off aggregation will also cause the primary provider to advertise all routes within the aggregate, and all benefits of aggregation will be lost.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) When a customer is connected to multiple providers, how can aggregation cause suboptimal path selection?

A) Aggregation causes all traffic to flow in the direction of the aggregated route.

B) Aggregation causes all attributes of the route to be lost, preventing neighboring BGP routers from completing path selection based on BGP attributes.

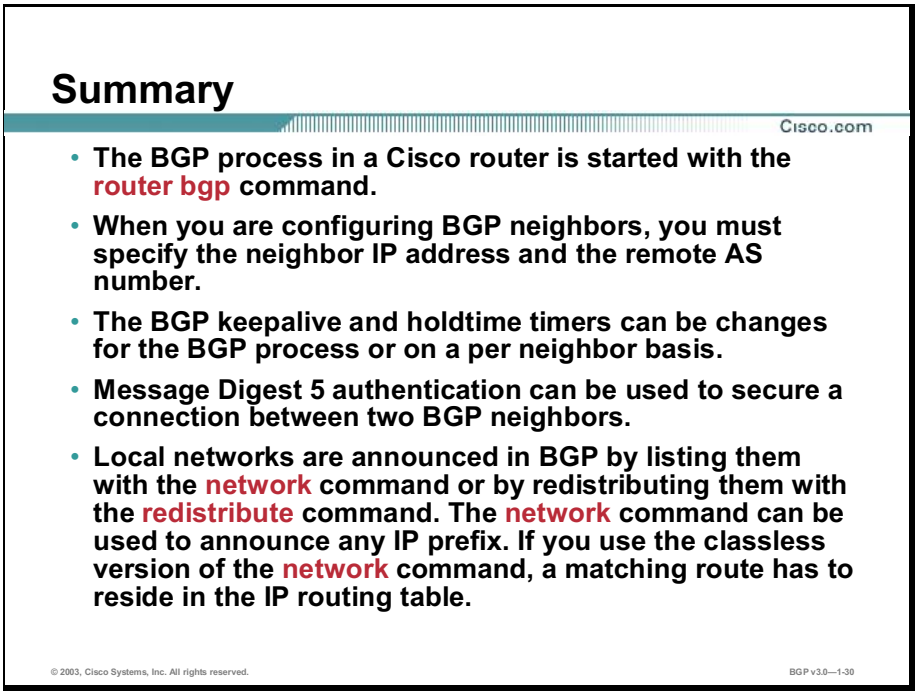
C) If both providers are not aggregating routes, the longest matching prefix rule will cause traffic to flow toward the nonaggregated route.

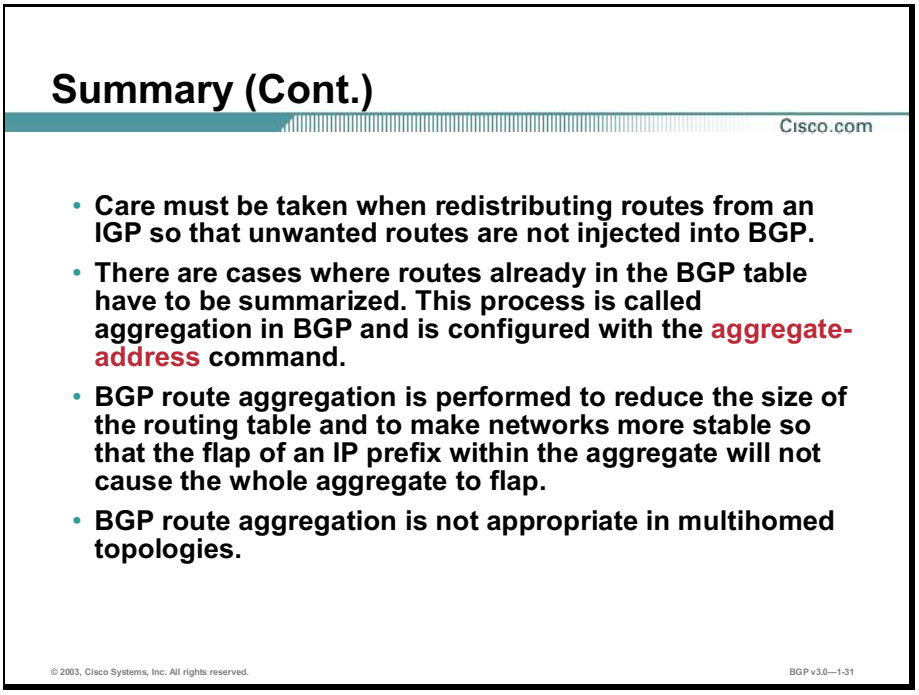
D) BGP will attempt to load-balance between multiple paths to the destination, forcing each session over a different network path.

Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• The BGP process in a Cisco router is started with the router bgp command.
• When you are configuring BGP neighbors, you must specify the neighbor IP address and the remote AS number.
• The BGP keepalive and holdtime timers can be changes for the BGP process or on a per neighbor basis.
• Message Digest 5 authentication can be used to secure a connection between two BGP neighbors.
• Local networks are announced in BGP by listing them with the network command or by redistributing them with the redistribute command. The network command can be used to announce any IP prefix. If you use the classless version of the network command, a matching route has to reside in the IP routing table.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-30


Summary (Cont.)
Cisco.com
• Care must be taken when redistributing routes from an IGP so that unwanted routes are not injected into BGP.
• There are cases where routes already in the BGP table have to be summarized. This process is called aggregation in BGP and is configured with the aggregate-address command.
• BGP route aggregation is performed to reduce the size of the routing table and to make networks more stable so that the flap of an IP prefix within the aggregate will not cause the whole aggregate to flap.
• BGP route aggregation is not appropriate in multihomed topologies.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-31

 Copyright © 2003, Cisco Systems, Inc.

Next Steps

After completing this lesson, go to:

- Monitoring and Troubleshooting BGP lesson

References

For additional information, refer to these resources:

- For more information on basic BGP configuration, refer to “Border Gateway Protocol” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

Quiz: Basic BGP Configuration

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Identify the Cisco IOS commands required to configure the BGP routing process
- Identify the Cisco IOS commands required to configure BGP neighbors
- Identify the Cisco IOS commands required to configure basic timers used in BGP
- Identify the Cisco IOS commands required to configure MD5 authentication for BGP
- Identify the commands required to announce local networks in BGP
- Describe BGP route redistribution and identify the commands required to configure BGP route redistribution
- Describe the classless behavior of BGP and identify the Cisco IOS commands required to configure BGP for classless operation
- Describe BGP route aggregation and identify the Cisco IOS commands required to configure basic BGP route aggregation
- Determine when BGP route aggregation is not appropriate in multihomed topologies

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What is the valid AS number range for a BGP process on a Cisco router?

A) 1 – 256

B) 1 – 32768

C) 1 – 65535

D) 1 – 131072

 Copyright © 2003, Cisco Systems, Inc.

Q2) What AS numbers are defined as private AS numbers?

A) 1 – 128

B) 32768 – 64511

C) 64512 – 65535

D) 65536 – 131072

Q3) What two parameters must you configure with the **neighbor** command to establish a BGP session with an external neighbor? (Choose two.)

A) neighbor IP address

B) subnet mask of the IP network

C) remote AS number

D) local AS number

E) description of the neighbor

Q4) What is the best method to temporarily disable a BGP neighbor session?

A) Remove the **neighbor** command from the BGP router process.

B) Remove the BGP router process from the configuration.

C) Terminate the neighbor connection with the **neighbor shutdown** command.

D) Disconnect the neighbor by initiating a router reload.

Q5) What three steps must you complete to advertise a classless prefix into BGP? (Choose three.)

A) Configure the prefix with the **network** command.

B) Specify the **mask** keyword with the locally advertised route.

C) Configure **redistribute connected** under the BGP router process.

D) Use a static route pointing to null 0 that matches the prefix.

Copyright © 2003, Cisco Systems, Inc.

Q6) What origin code is carried with routes redistributed into BGP?

A) internal

B) external

C) unknown

D) incomplete

Q7) What must be true for a BGP route aggregate to be advertised in the IP routing table?

A) At least one network in the specified range must exist in the BGP table.

B) You must configure automatic summarization under the BGP routing protocol process.

C) You must configure a route to null 0 matching the aggregate.

D) No synchronization must be configured under the BGP routing protocol process.

Q8) What are two benefits of using route aggregation in BGP? (Choose two.)

A) It ensures that even if aggregate networks are down, the aggregate is advertised, eliminating “black holes.”

B) It reduces the amount of memory used in the router to store the BGP table.

C) It reduces route flapping and its effects on router CPU resources.

D) BGP attribute granularity is maintained, ensuring optimal path selection.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring and Troubleshooting BGP

Overview

This lesson introduces the different Cisco IOS® commands available for monitoring and troubleshooting basic Border Gateway Protocol (BGP) configurations. The commands required to monitor the status of BGP, neighbor connections, and the BGP table are discussed. The lesson also discusses techniques for troubleshooting the most common BGP session startup issues.

Importance

BGP monitoring commands are important to ensuring that basic BGP configurations are operating correctly. If basic BGP configurations are not functioning as expected, BGP troubleshooting skills are critical to successful problem resolution.

Objectives

Upon completing this lesson, you will be able to:

- Identify the Cisco IOS commands required to monitor the status of the BGP routing process
- Identify the Cisco IOS commands required to monitor BGP neighbors
- Identify the Cisco IOS commands required to monitor the BGP table
- Identify the Cisco IOS commands required to perform basic BGP debugging
- List common BGP session startup problems
- Troubleshoot basic BGP session startup problems when the neighbor is not reachable

- Troubleshoot basic BGP session startup problems when the neighbor is not configured
- Troubleshoot basic BGP session startup problems when an AS number mismatch exists

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Successful completion of *Building Scalable Cisco Internetworks* (BSCI) or equivalent

Outline

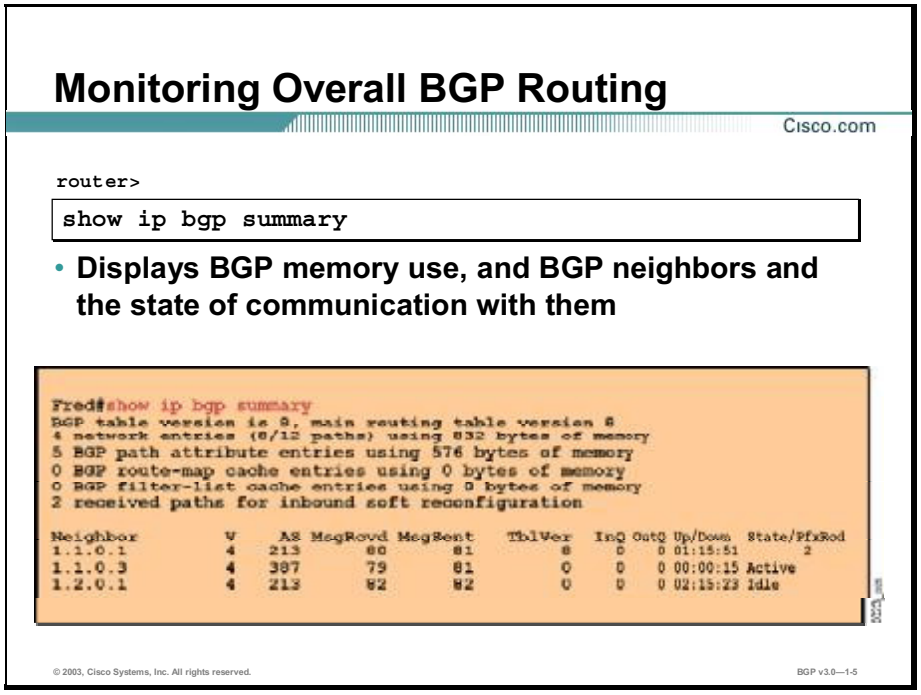
This lesson includes these topics:

- Overview
- Monitoring Overall BGP Routing
- Monitoring BGP Neighbors
- Monitoring the BGP Table
- Debugging BGP
- BGP Session Startup Problems
- BGP Neighbor Not Reachable
- BGP Neighbor Not Configured
- BGP AS Number Mismatch
- Summary
- Assessment (Lab): BGP Overview

 Copyright © 2003, Cisco Systems, Inc.

Monitoring Overall BGP Routing

This topic presents the command used to monitor the overall status of the BGP routing protocol process.


Monitoring Overall BGP Routing
Cisco.com
router>
show ip bgp summary
• Displays BGP memory use, and BGP neighbors and the state of communication with them
Fred#show ip bgp summary
BGP table version is 8, main routing table version 8
4 network entries (8/12 paths) using 832 bytes of memory
5 BGP path attribute entries using 576 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
2 received paths for inbound soft reconfiguration
Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd
1.1.0.1 4 213 60 61 8 0 0 01:15:51 2
1.1.0.3 4 307 79 61 0 0 0 00:00:15 Active
1.2.0.1 4 213 62 62 0 0 0 02:15:23 Idle
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-5

This is a very useful command when you are troubleshooting BGP. The output provides a short summary of the status of the BGP process in the router.

The first section in the output here describes the BGP table and its content:

- The “BGP table version” is the version number of the local BGP table. This number is incremented every time that the table is changed.
- The “main routing table version” shows the last version of the BGP database that was injected into the main routing table.
- The subsequent lines of text indicate the amount of memory allocated to hold the table. The output displays how many networks are known and how many different paths and attribute values are associated with them.

The second section of the output is a table in which the current neighbor statuses are shown. There is one line of text for each neighbor configured. The columns are:

- IP address of the neighbor as configured in the local router.
- BGP version number used by the router when communicating with the neighbor.
- Autonomous system (AS) number of the remote neighbor.

- Number of messages/updates received from the neighbor since the session was established.
- Number of messages/updates sent to the neighbor since the session was established.
- Version number of the local BGP table that has been included in the most recent update to the neighbor.
- Number of messages waiting to be processed in the incoming queue from this neighbor.
- Number of messages waiting in the outgoing queue for transmission to the neighbor.
- How long the neighbor has been in the current state and the name of the current state. State "Established" is not printed out, so no state name indicates the state "Established".

You can use this information to verify that BGP sessions are up and established. If they are not, you will have to further investigate the BGP configuration to locate the problem. You can also verify the IP address and AS number of the configured BGP neighbor with the **show ip bgp summary** command.

If the session is "Established", the number of messages sent and received, as displayed in the output of the **show ip bgp summary** command, can indicate BGP stability. Use the command a few times, with a time interval between the printouts, and calculate how many messages have been exchanged during that period.

A large number of messages in the in queue indicates a lack of CPU resources in the local router. A large number of messages in the out queue indicates a lack of bandwidth to the remote router or a lack of CPU resources in the remote router.

show ip bgp summary

To display the status of all BGP connections, use the **show ip bgp summary** EXEC command.

show ip bgp summary

Syntax Description

This command has no arguments or keywords.

Practice

Q1) What three key pieces of information are contained in the output of the **show ip bgp summary** command? (Choose three.)

A) BGP memory use

B) BGP neighbors

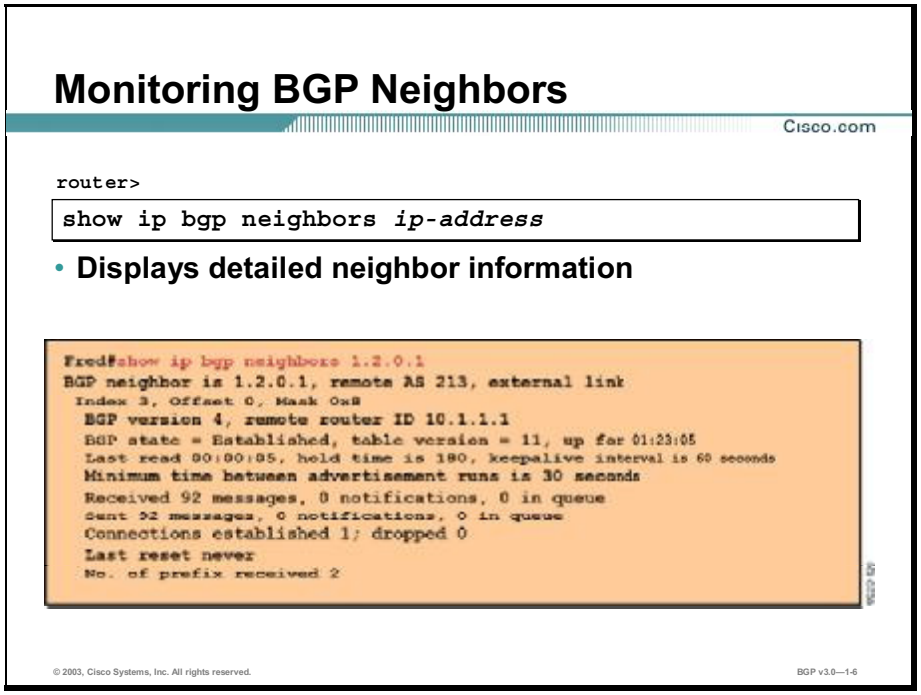
C) BGP route table

D) BGP neighbor connection state

 Copyright © 2003, Cisco Systems, Inc.

Monitoring BGP Neighbors

This topic presents the Cisco IOS command used to monitor BGP neighbors.


Monitoring BGP Neighbors
Cisco.com
router>
show ip bgp neighbors ip-address
• Displays detailed neighbor information
Fred#show ip bgp neighbors 1.2.0.1
BGP neighbor is 1.2.0.1, remote AS 213, external link
Index 3, Offset 0, Mask 0x8
BGP version 4, remote router ID 10.1.1.1
BGP state = Established, table version = 11, up for 01:23:05
Last read 00:00:05, hold time is 180, keepalive interval is 60 seconds
Minimum time between advertisement runs is 30 seconds
Received 92 messages, 0 notifications, 0 in queue
Sent 92 messages, 0 notifications, 0 in queue
Connections established 1; dropped 0
Last reset never
No. of prefix received 2
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-6

You can use this command for two different purposes. The general purpose, as shown in the figure, is to get information about the TCP session and the BGP parameters of the session. All BGP session parameters are displayed. In addition, TCP timers and counters are also displayed.

The other use is not shown in this example. If any of the optional qualifiers referring to routes or paths are given, the BGP routing information sent or received on this session will be displayed. This feature is useful when you are troubleshooting path selection.

show ip bgp neighbors

To display information about the TCP and BGP connections to neighbors, use the **show ip bgp neighbors** EXEC command.

show ip bgp neighbors [*address*] [**received-routes** | **routes** | **advertised-routes** | {**paths** *regular-expression*} | **dampened-routes**]

Syntax Description

address	(Optional) Address of the neighbor whose routes you have learned from. If you omit this argument, all neighbors are displayed.
received-routes	(Optional) Displays all received routes (both accepted and rejected) from the specified neighbor.
routes	(Optional) Displays all routes that are received and accepted. This is a subset of the output from the **received-routes** keyword.

advertised-routes (Optional) Displays all the routes that the router has advertised to the neighbor.

paths *regular-expression*
(Optional) Regular expression that is used to match the paths received.

dampened-routes (Optional) Displays the dampened routes to the neighbor at the IP address specified.

Practice

Q1) Which command do you use to display detailed BGP neighbor information?

A) **show ip bgp summary**

B) **show ip bgp**

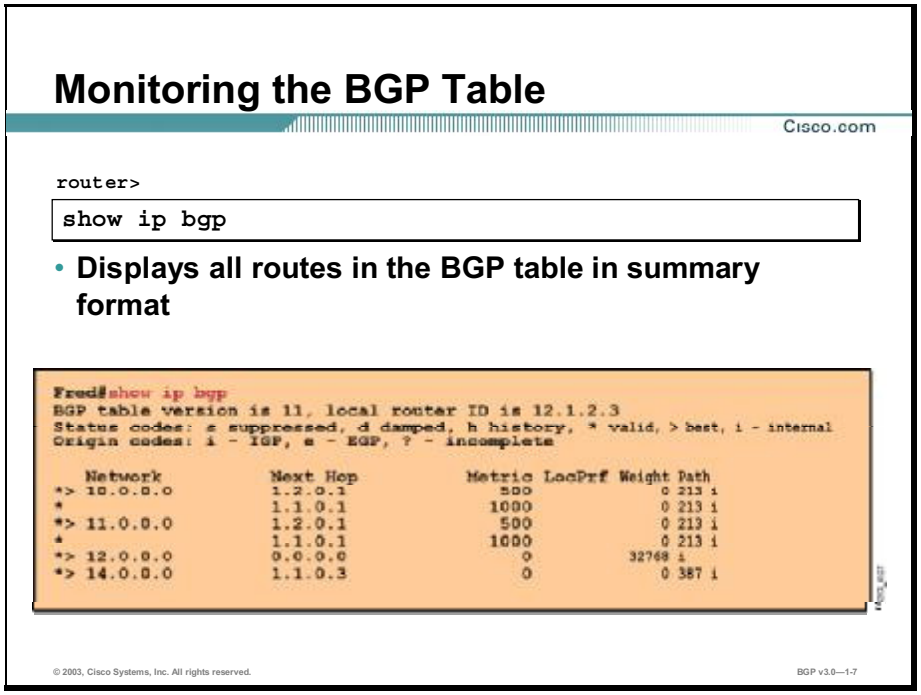
C) **show ip bgp neighbors** *address*

D) **show ip bgp detail**

 Copyright © 2003, Cisco Systems, Inc.

Monitoring the BGP Table

This topic lists the Cisco IOS commands used to monitor the BGP routing table.


Monitoring the BGP Table
Cisco.com
router>
show ip bgp
• Displays all routes in the BGP table in summary format
Fred#show ip bgp
BGP table version is 11, local router ID is 12.1.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 10.0.0.0 1.2.0.1 500 0 213 i
* 1.1.0.1 1000 0 213 i
*> 11.0.0.0 1.2.0.1 500 0 213 i
* 1.1.0.1 1000 0 213 i
*> 12.0.0.0 0.0.0.0 0 32768 i
*> 14.0.0.0 1.1.0.3 0 0 387 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-7

In most cases, when the **show ip bgp** command is given without optional qualifiers, the entire BGP table is displayed. An abbreviated list of information about each route is displayed, one line per prefix. The output is sorted in network number order. Therefore, if the BGP table contains more than one route to the same network, both routes are displayed on successive lines. The network number is printed on the first of those lines only. The following lines, which are referring to the same network, have the network number field left blank.

Some, but not all, of the BGP attributes associated with the route are displayed on the line. Next hop, multi-exit discriminator (MED; displayed as “Metric”), local preference, and weight each have their own columns. The AS-path attribute is displayed as the sequence of AS numbers in the Path column. Immediately following the AS path, but not part of the AS-path attribute, the origin attribute is displayed. The lowercase letter “i” means origin code Interior Gateway Protocol(IGP); “e” means exterior gateway protocol (EGP) and ‘?’ means incomplete/unknown.

The BGP path selection will select one of the alternative routes to each of the networks as the best. This route will be pointed out by the character “>” in the left column.

show ip bgp

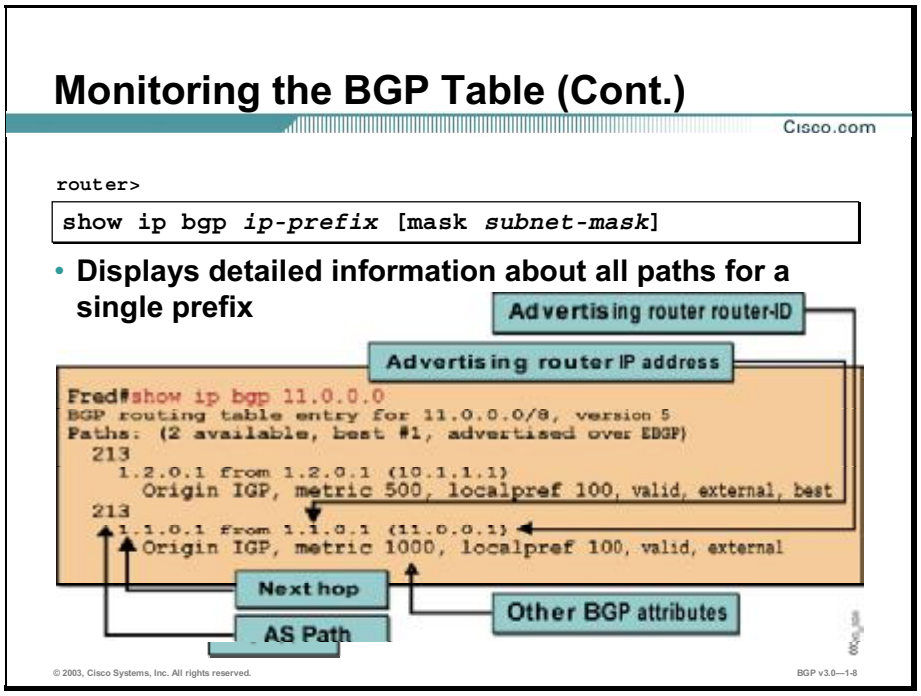
To display entries in the BGP routing table, use the **show ip bgp** EXEC command.

show ip bgp [*network*] [*network-mask*] [**longer-prefixes**]

Syntax Description

network	(Optional) Network number, entered to display a particular network in the BGP routing table
network-mask	(Optional) Displays all BGP routes matching the address/mask pair
longer-prefixes	(Optional) Displays route and more specific routes

 Copyright © 2003, Cisco Systems, Inc.


Monitoring the BGP Table (Cont.)
Cisco.com
router>
show ip bgp ip-prefix [mask subnet-mask]
• Displays detailed information about all paths for a single prefix
Advertising router router-ID
Advertising router IP address
Fred#show ip bgp 11.0.0.0
BGP routing table entry for 11.0.0.0/8, version 5
Paths: (2 available, best #1, advertised over EBGP)
213
1.2.0.1 from 1.2.0.1 (10.1.1.1)
Origin IGP, metric 500, localpref 100, valid, external, best
213
1.1.0.1 from 1.1.0.1 (11.0.0.1)
Origin IGP, metric 1000, localpref 100, valid, external
Next hop
AS Path
Other BGP attributes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-8

If more information and the complete set of BGP attributes are required, the **show ip bgp** command should be given with the network number on the command line. This command displays all relevant BGP information about that specific network.

In this example, the information about network 11.0.0.0 is displayed. There are two different routes to 11.0.0.0. One is received from neighbor 1.2.0.1 and the other from 1.1.0.1.

The BGP route selection process has selected the route via 1.2.0.1 as the best. This is thus the route that BGP will try to install in the forwarding table. Installation of routes in the forwarding table is made based on the administrative distance (AD).

Practice

Q1) How does the output from the **show ip bgp** command indicate which route to a specific destination is selected as the best?

A) The route selected as the best is enclosed in parentheses: “( )”.

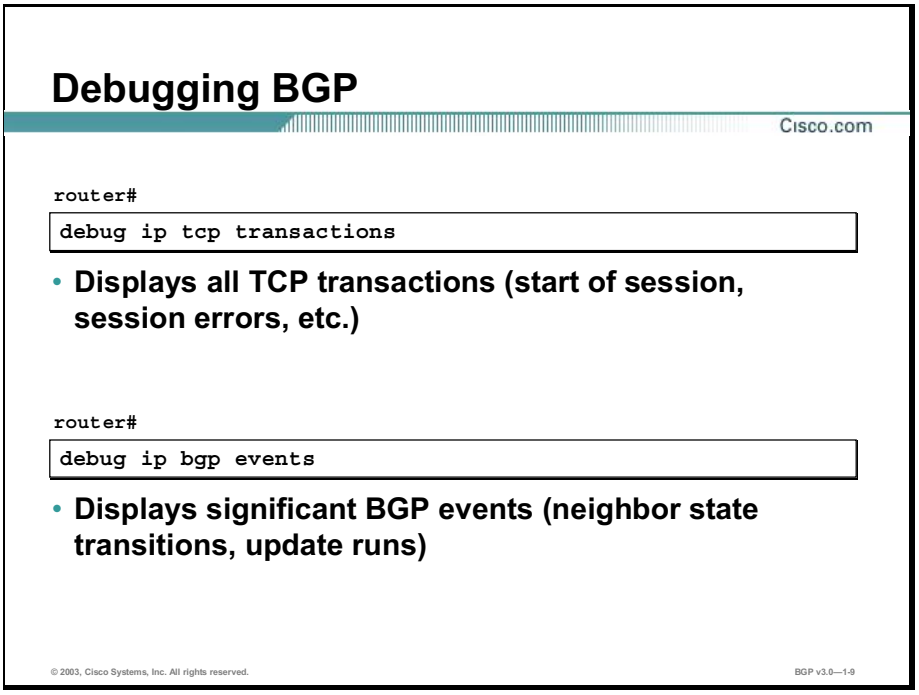
B) The route selected as the best is marked with the character “>”.

C) The route selected as best is marked with an asterisk: “*”.

D) Only the best route is inserted into the BGP table.

Debugging BGP

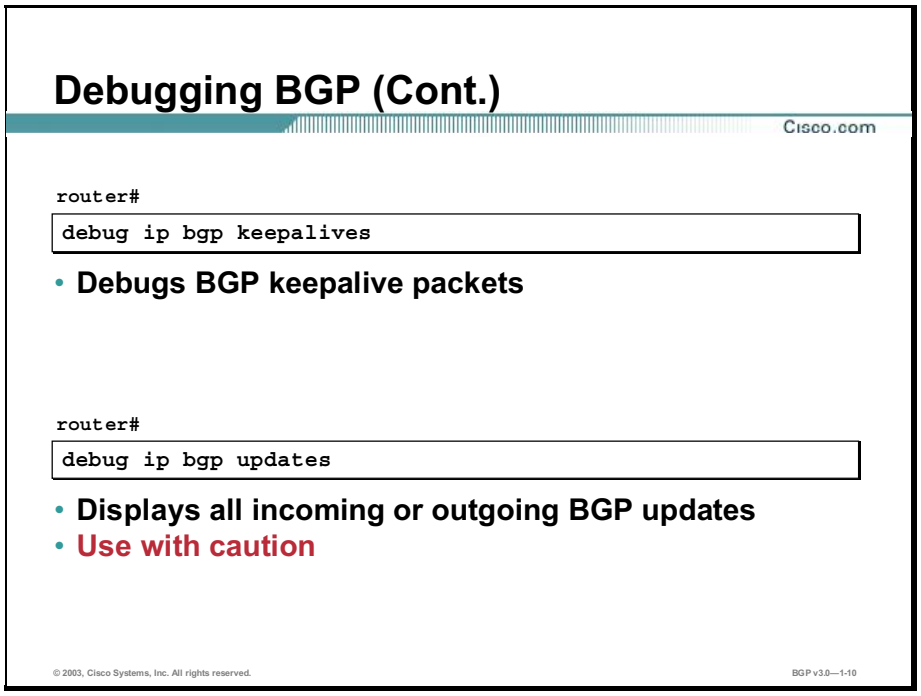
This topic lists the Cisco IOS commands used to perform debugging of basic BGP configurations.


Debugging BGP
Cisco.com
router#
debug ip tcp transactions
• Displays all TCP transactions (start of session, session errors, etc.)
router#
debug ip bgp events
• Displays significant BGP events (neighbor state transitions, update runs)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-9

If a BGP session stays in the Active state, where it is actively sending connection attempts to the neighbor, **debug ip tcp transactions** can give valuable information about failed connection attempts. All TCP transactions in the router are displayed on the console as they happen. The network administrator can now determine if the TCP session is being established, and, if not, what the probable cause might be.

If the TCP session succeeds, but is torn down within a short period of time, the reason might be found using **debug ip bgp events**. All BGP events will be displayed on the console as they happen if this debug command is enabled.

 Copyright © 2003, Cisco Systems, Inc.



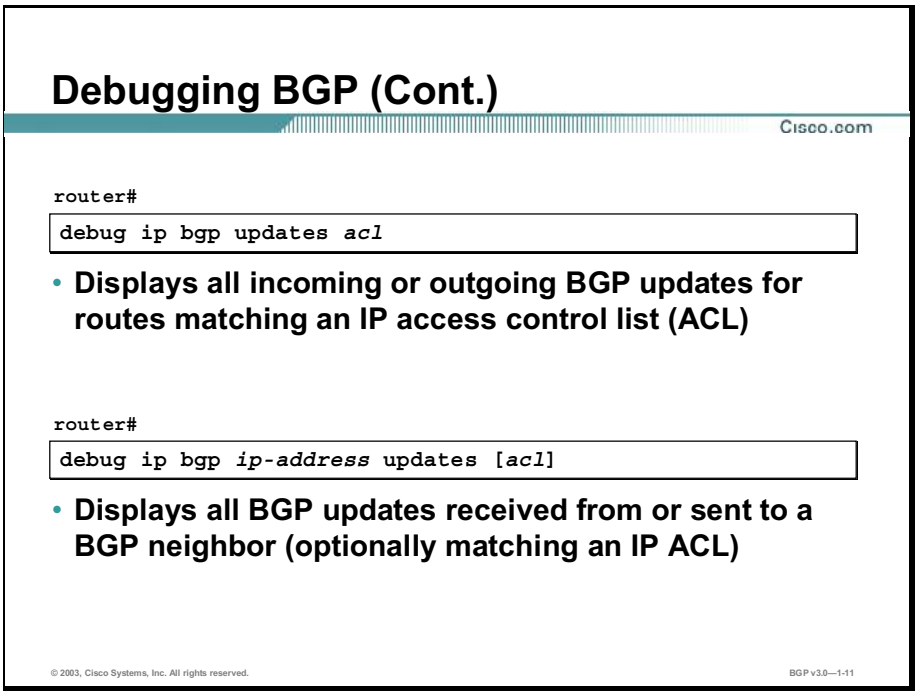
In a stable state with no network topology changes, no BGP updates are sent between neighboring routers. When a BGP session has been idle for some time, the BGP protocol will exchange keepalive packets between BGP neighbors. The keepalive timer has a default value of 60 seconds.

Use **debug ip bgp keepalives** to get a printout on the console for every keepalive packet sent or received. Successful keepalive exchanges indicate that the session is working and is in a stable state.

If no keepalives are sent or received, the session might still be working. The reason for not seeing any keepalives is that the session is never idle long enough.

Use **debug ip bgp updates** to get a printout on the console for every update message sent or received. The successful exchange of updates indicates that the session is working and is not “Idle”.

In a large network, updates are sent and received in large volumes. Starting **debug ip bgp updates** might cause extensive output on the console. In some cases, the CPU resources used to generate those outputs are so great that the real work that must be done will suffer. In a case with very busy BGP sessions, it is actually possible to set the router in a condition where all CPU resources are consumed with the debugging printouts.


Debugging BGP (Cont.)
Cisco.com
router#
debug ip bgp updates acl
• Displays all incoming or outgoing BGP updates for routes matching an IP access control list (ACL)
router#
debug ip bgp ip-address updates [acl]
• Displays all BGP updates received from or sent to a BGP neighbor (optionally matching an IP ACL)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-11

To avoid debug printouts for every update sent or received, you can create and associate an access-list with the **debug** command. When you use this command, the console will display only the updates referring to a network number permitted by the access-list. The command is extremely useful in a live network with busy BGP sessions where the troubleshooter is interested only in updates for specific networks.

Indicating a specific neighbor can even further restrict the debugging. The console will display only the updates on the session to the indicated neighbor. Optionally, you can combine this debug command with an access-list.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What debug command should you enable to troubleshoot BGP session startup issues where the TCP connection never succeeds?

A) **ip bgp updates**

B) **ip packets**

C) **ip bgp keepalives**

D) **ip tcp transactions**

Q2) What two steps can you take to limit the amount of debug output when debugging BGP updates? (Choose two.)

A) specify the incoming interface

B) limit the debug to a specific BGP neighbor

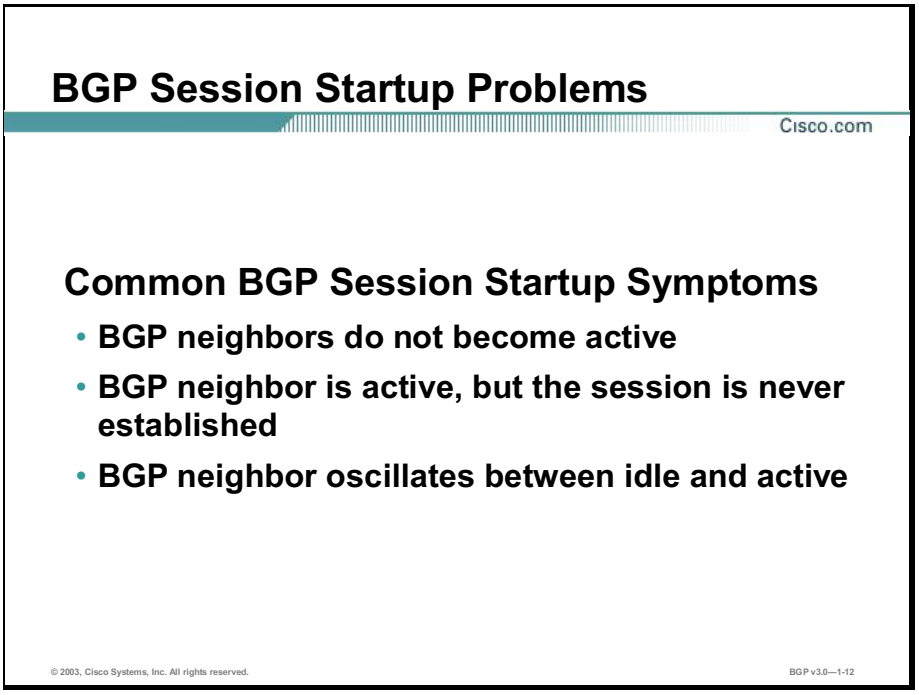
C) limit the debug to specific networks using ACLs

D) increase the keepalive timer to reduce BGP update traffic

Copyright © 2003, Cisco Systems, Inc.

BGP Session Startup Problems

This topic lists the most common session startup issues you can experience when configuring basic BGP.


BGP Session Startup Problems
Cisco.com
Common BGP Session Startup Symptoms
• BGP neighbors do not become active
• BGP neighbor is active, but the session is never established
• BGP neighbor oscillates between idle and active
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-12

There are a number of common BGP session startup symptoms:

- A BGP neighbor never becomes active.
- A BGP neighbor is active, but the BGP session is not established.
- The BGP neighbor state oscillates between idle and active.

Practice

Q1) What are three common BGP session startup issues? (Choose three.)

A) The BGP neighbor never becomes active.

B) The BGP neighbor state oscillates between idle and active.

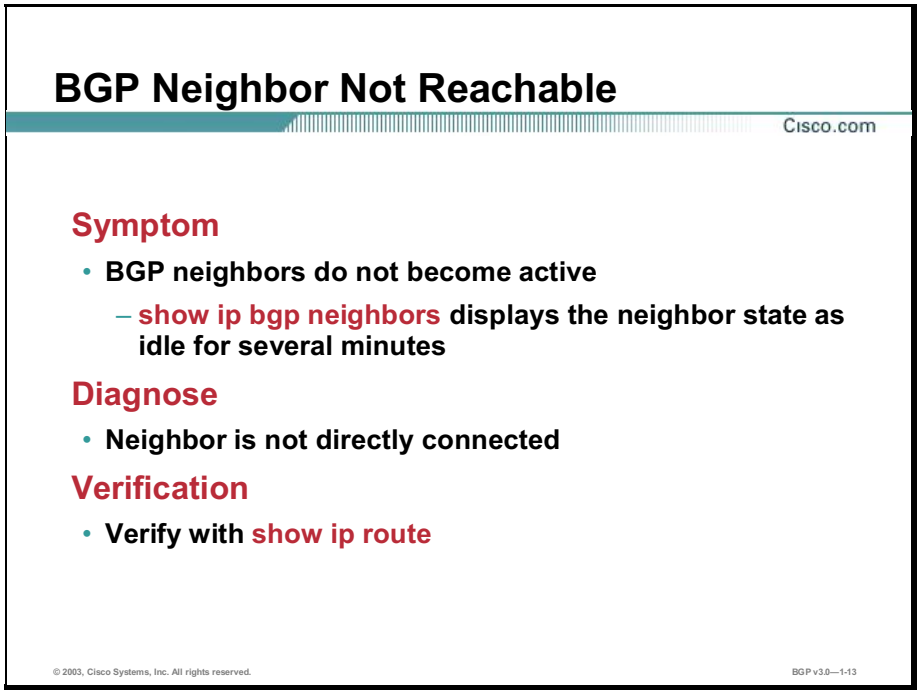
C) The BGP session is established, but the neighbor state is idle.

D) The BGP neighbor is active, but the BGP session is not established.

 Copyright © 2003, Cisco Systems, Inc.

BGP Neighbor Not Reachable

This topic describes basic BGP troubleshooting for BGP session startup problems where the neighbor is not reachable.

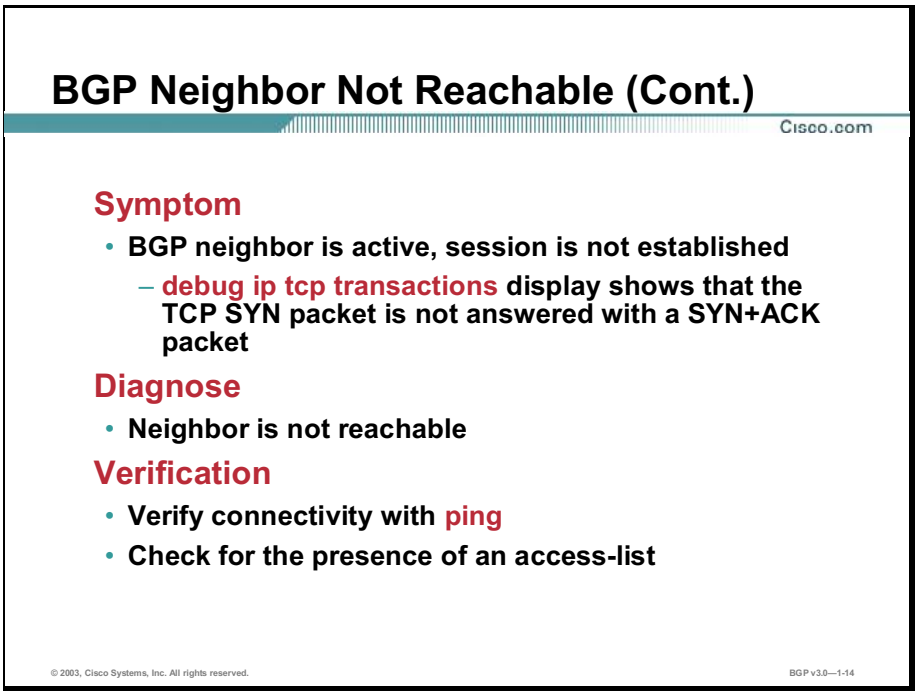

BGP Neighbor Not Reachable
Cisco.com
Symptom
• BGP neighbors do not become active
– show ip bgp neighbors displays the neighbor state as idle for several minutes
Diagnose
• Neighbor is not directly connected
Verification
• Verify with show ip route
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-13

BGP sessions to a router in another AS should normally run across directly connected interfaces (routers that share a common IP subnet). You must configure neighboring routers to reach each other using the IP address belonging to this shared subnet, so that no other routing protocol is required to set up the BGP session.

If a router is configured with a BGP neighbor that is in another AS but not directly connected, the session will stay in the Idle state. The router will not even attempt to set up the session.

The normal way to fix this problem is to change the neighbor reference so that it is referred by an IP address that is directly connected. However, in some odd cases, the neighbor is intentionally reachable using an interface that is not directly connected. In that rare case, the local router must have routing information on how to reach that address. Also, you must configure the BGP session with the **ebgp-multihop** option.

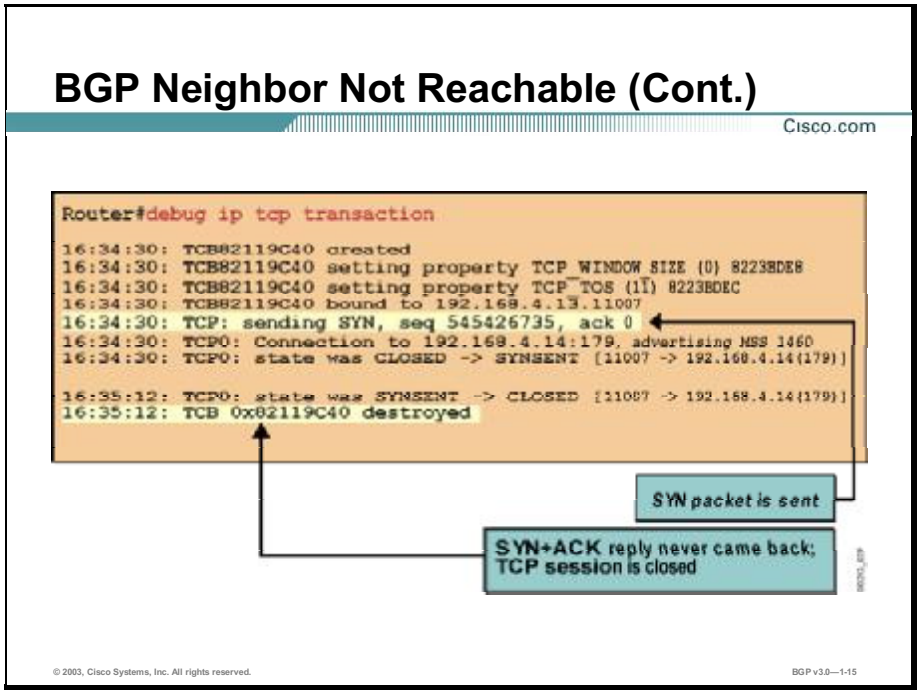
If the session goes into the Active state, the router will attempt to establish the session. If session establishment is unsuccessful, you will have to troubleshoot the problem. The **debug ip tcp transactions** command will display the connect attempts.


BGP Neighbor Not Reachable (Cont.)
Cisco.com
Symptom
• BGP neighbor is active, session is not established
– debug ip tcp transactions display shows that the TCP SYN packet is not answered with a SYN+ACK packet
Diagnose
• Neighbor is not reachable
Verification
• Verify connectivity with ping
• Check for the presence of an access-list
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-14

TCP session establishment starts with the router sending a TCP SYN packet. If the TCP SYN packet is never answered, the remote router might be dead or not reachable. Try to use the **ping** command and verify the existence of the remote router and the IP packet exchange between the local and remote router.

 Copyright © 2003, Cisco Systems, Inc.

Example


BGP Neighbor Not Reachable (Cont.)
Cisco.com
Router#debug ip tcp transaction
16:34:30: TCB82119C40 created
16:34:30: TCB82119C40 setting property TCP_WINDOW_SIZE (0) 8223BDE8
16:34:30: TCB82119C40 setting property TCP_TOS (11) 8223BDEC
16:34:30: TCB82119C40 bound to 192.168.4.13.11007
16:34:30: TCP: sending SYN, seq 545426735, ack 0
16:34:30: TCP0: Connection to 192.168.4.14:179, advertising MSS 1460
16:34:30: TCP0: state was CLOSED -> SYNSENT [11007 -> 192.168.4.14(179)]
16:35:12: TCP0: state was SYNSENT -> CLOSED [11007 -> 192.168.4.14(179)]
16:35:12: TCB 0x82119C40 destroyed
SYN packet is sent
SYN+ACK reply never came back; TCP session is closed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-15

In the scenario shown here, the remote BGP router is not available. The sending router therefore never receives the reply to the SYN packet and aborts the TCP session in approximately 45 seconds (changing the state from synsent to closed).

Practice

Q1) What is the most common reason for a BGP session not leaving the Idle state?

A) The TCP port for the connection is not configured.

B) The external neighbor is not directly connected.

C) The TCP SYN packet is answered with an RST packet.

D) The neighbors have been configured with the same AS number.

Q2) What can you use to establish a BGP session to external neighbors that are not directly connected?

A) IBGP

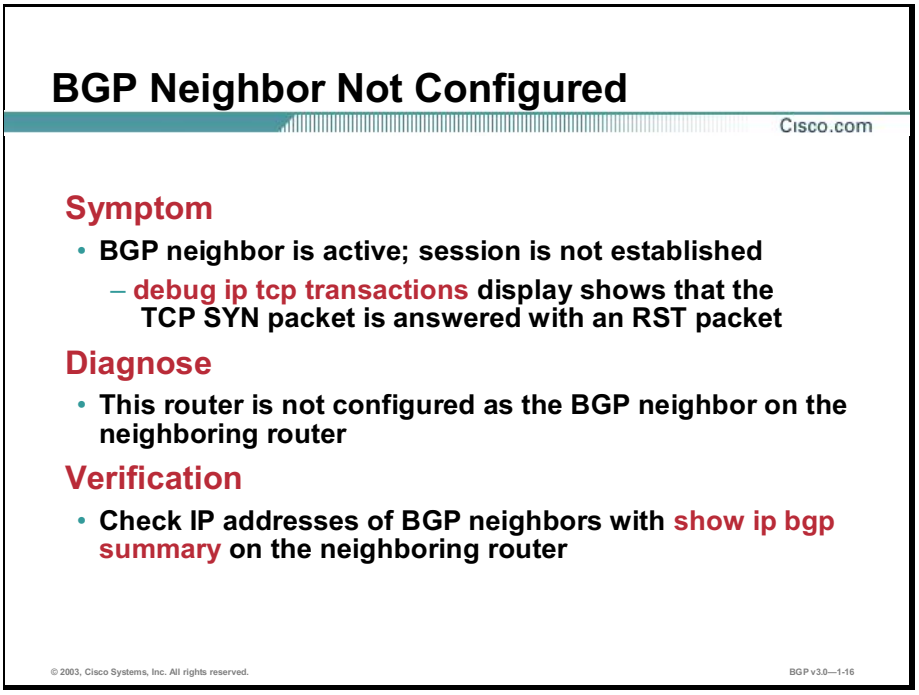
B) an IGP

C) static routes

D) **ebgp-multihop**

BGP Neighbor Not Configured

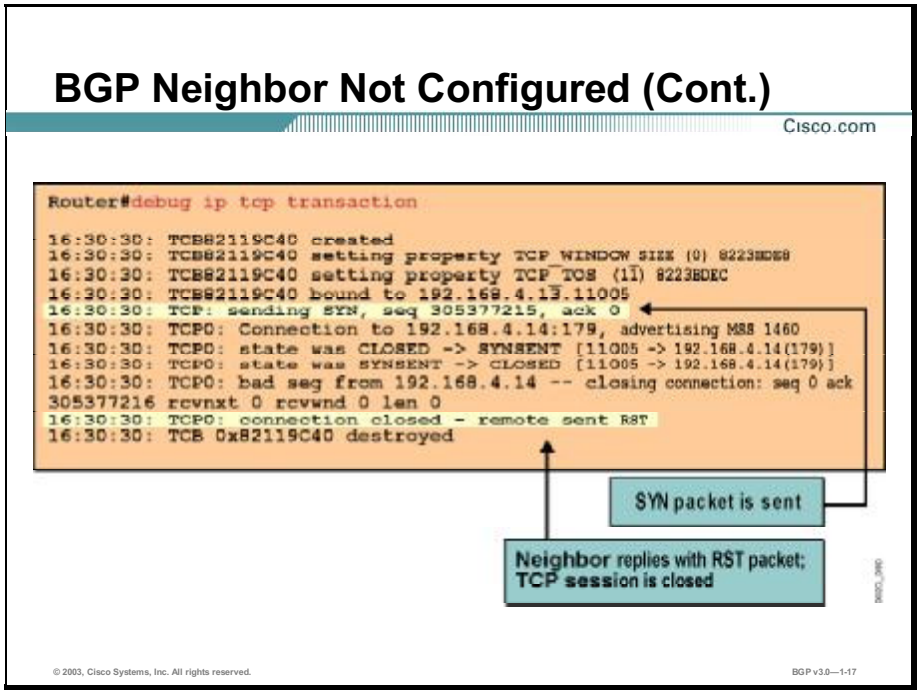
This topic describes basic BGP troubleshooting for BGP session startup problems where the neighbor is not configured.


BGP Neighbor Not Configured
Cisco.com
Symptom
• BGP neighbor is active; session is not established
– debug ip tcp transactions display shows that the TCP SYN packet is answered with an RST packet
Diagnose
• This router is not configured as the BGP neighbor on the neighboring router
Verification
• Check IP addresses of BGP neighbors with show ip bgp summary on the neighboring router
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-16

If the TCP SYN packet is answered with a TCP RST packet, the remote router is alive and reachable but is not willing to grant the connection attempt. The reason for this refusal may be that BGP is not started on the remote router or that the source IP address used by the local router in the connection attempt is not in the list of valid neighbors for the remote router.

 Copyright © 2003, Cisco Systems, Inc.

Example


BGP Neighbor Not Configured (Cont.)
Cisco.com
Router#debug ip tcp transaction
16:30:30: TCB82119C40 created
16:30:30: TCB82119C40 setting property TCP WINDOW SIZE (0) 8223BDE8
16:30:30: TCB82119C40 setting property TCP TOS (11) 8223BDEC
16:30:30: TCB82119C40 bound to 192.168.4.13.11005
16:30:30: TCP: sending SYN, seq 305377215, ack 0
16:30:30: TCP0: Connection to 192.168.4.14:179, advertising MSS 1460
16:30:30: TCP0: state was CLOSED -> SYNSENT [11005 -> 192.168.4.14(179)]
16:30:30: TCP0: state was SYNSENT -> CLOSED [11005 -> 192.168.4.14(179)]
16:30:30: TCP0: bad seg from 192.168.4.14 -- closing connection: seg 0 ack
305377216 rcvnxt 0 rcvwnd 0 len 0
16:30:30: TCP0: connection closed - remote sent RST
16:30:30: TCB 0x82119C40 destroyed
SYN packet is sent
Neighbor replies with RST packet; TCP session is closed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-17

In the scenario shown here, the remote router is not configured for BGP or there was a mismatch in the neighbor IP addresses. The remote router responds with an RST packet as soon as it receives the initial SYN packet, terminating the BGP session.

Practice

Q1) What will result from attempting to open a BGP connection with a neighbor that has not been properly configured for BGP?

A) The BGP session will remain in the Idle state.

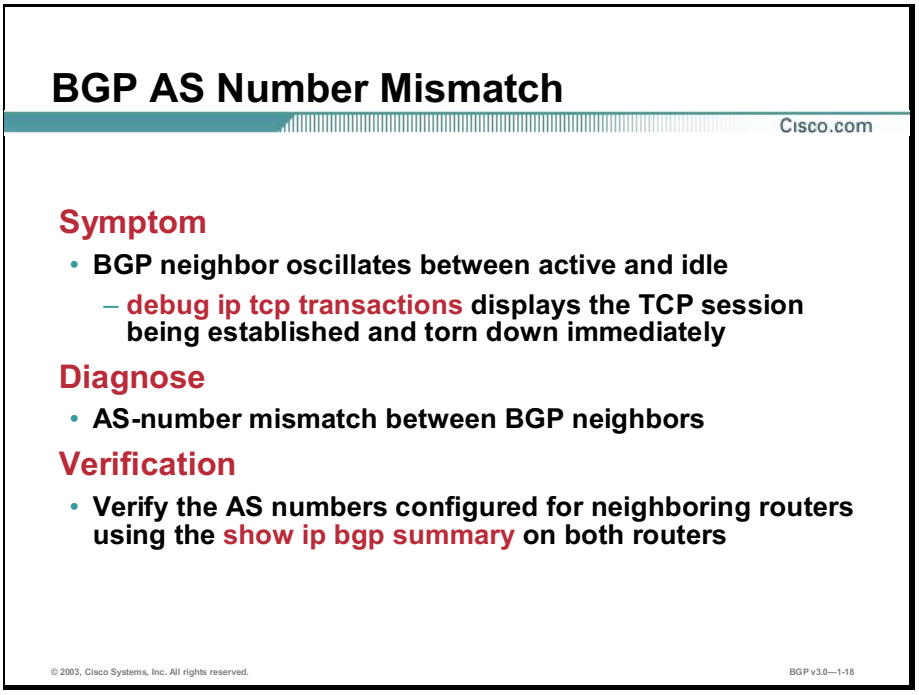
B) The neighbor session will be established and the session startup parameters will be negotiated over the TCP session.

C) The BGP session is immediately terminated with a TCP RST packet.

D) The BGP session will become “stuck in Active state.”

BGP AS Number Mismatch

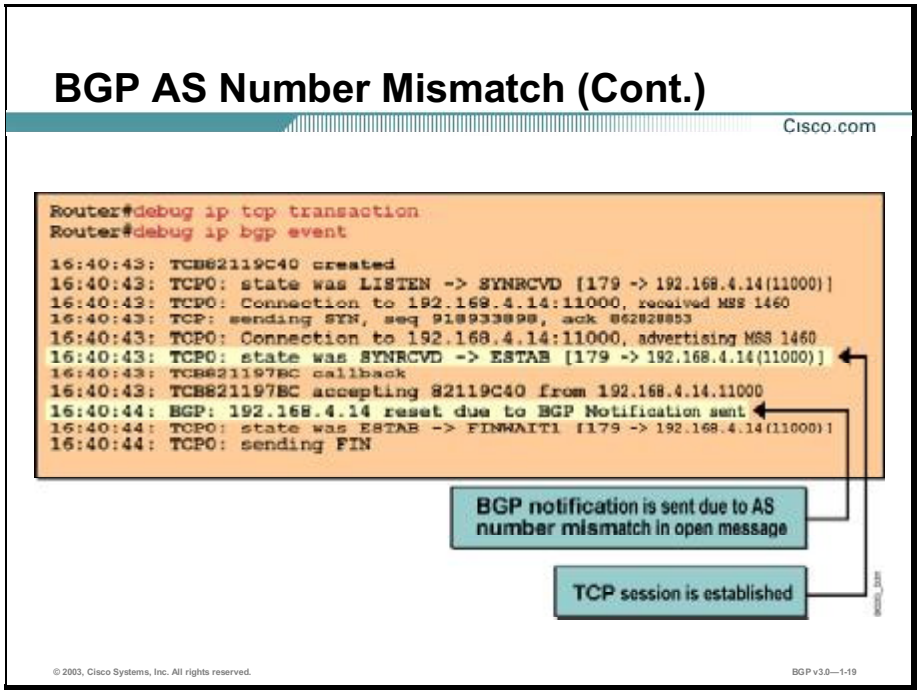
This topic describes basic BGP troubleshooting for BGP session startup problems where the AS numbers are not properly configured.


BGP AS Number Mismatch
Cisco.com
Symptom
• BGP neighbor oscillates between active and idle
– debug ip tcp transactions displays the TCP session being established and torn down immediately
Diagnose
• AS-number mismatch between BGP neighbors
Verification
• Verify the AS numbers configured for neighboring routers using the show ip bgp summary on both routers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-18

If the TCP session is established using the specified three-way handshake, SYN, SYN-ACK, ACK, but the router drops the session after a short packet exchange, the BGP parameters are mismatched. Make sure that the remote AS configured on each router matches the local AS configured on the neighbor. If the AS numbers do not match, the router will drop the session after exchanging BGP Open messages.

 Copyright © 2003, Cisco Systems, Inc.

Example


BGP AS Number Mismatch (Cont.)
Cisco.com
Router#debug ip tcp transaction
Router#debug ip bgp event
16:40:43: TCB82119C40 created
16:40:43: TCP0: state was LISTEN -> SYNRCVD [179 -> 192.168.4.14(11000)]
16:40:43: TCP0: Connection to 192.168.4.14:11000, received MSS 1460
16:40:43: TCP: sending SYN, seq 918933898, ack 862828853
16:40:43: TCP0: Connection to 192.168.4.14:11000, advertising MSS 1460
16:40:43: TCP0: state was SYNRCVD -> ESTAB [179 -> 192.168.4.14(11000)]
16:40:43: TCB821197BC callback
16:40:43: TCB821197BC accepting 82119C40 from 192.168.4.14.11000
16:40:44: BGP: 192.168.4.14 reset due to BGP Notification sent
16:40:44: TCP0: state was ESTAB -> FINWAIT1 [179 -> 192.168.4.14(11000)]
16:40:44: TCP0: sending FIN
BGP notification is sent due to AS number mismatch in open message
TCP session is established
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-19

Whenever there is a mismatch in AS numbers (or any other BGP parameters that are necessary for proper BGP operation), the BGP session is terminated with a BGP notification, and the TCP session is terminated as well.

Practice

Q1) What symptom usually indicates an AS number mismatch?

A) The BGP session is established and immediately closed.

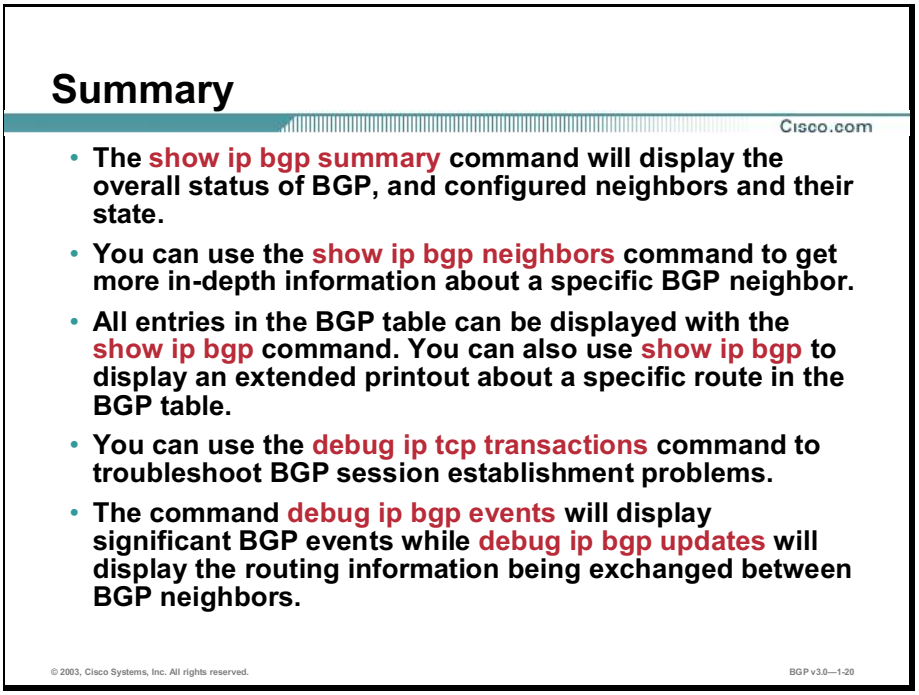
B) A TCP RST packet is sent.

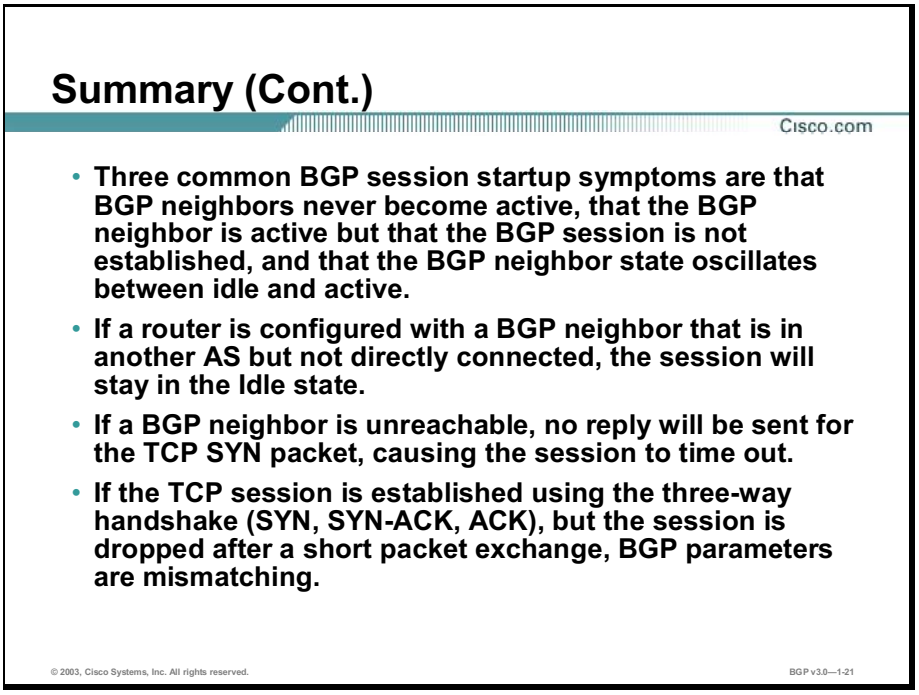
C) The TCP session goes into synsent state.

D) The BGP session remains in Idle state.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• The show ip bgp summary command will display the overall status of BGP, and configured neighbors and their state.
• You can use the show ip bgp neighbors command to get more in-depth information about a specific BGP neighbor.
• All entries in the BGP table can be displayed with the show ip bgp command. You can also use show ip bgp to display an extended printout about a specific route in the BGP table.
• You can use the debug ip tcp transactions command to troubleshoot BGP session establishment problems.
• The command debug ip bgp events will display significant BGP events while debug ip bgp updates will display the routing information being exchanged between BGP neighbors.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-20


Summary (Cont.)
Cisco.com
• Three common BGP session startup symptoms are that BGP neighbors never become active, that the BGP neighbor is active but that the BGP session is not established, and that the BGP neighbor state oscillates between idle and active.
• If a router is configured with a BGP neighbor that is in another AS but not directly connected, the session will stay in the Idle state.
• If a BGP neighbor is unreachable, no reply will be sent for the TCP SYN packet, causing the session to time out.
• If the TCP session is established using the three-way handshake (SYN, SYN-ACK, ACK), but the session is dropped after a short packet exchange, BGP parameters are mismatching.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—1-21

 Copyright © 2003, Cisco Systems, Inc.

Next Steps

After completing this lesson, go to:

- Route Selection Using Policy Controls module

References

For additional information, refer to these resources:

- For more information on BGP monitoring and troubleshooting, refer to “Border Gateway Protocol” at the following URL:
  http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

- For further information on BGP monitoring and troubleshooting, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL:
  http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: BGP Overview (Initial Lab Setup)

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will perform initial router configuration, configure the interfaces on your routers, and establish IGP connectivity across your core backbone.

After completing this exercise, you will be able to:

- Use your BGP student workgroup to complete the BGP lab exercises in the *Configuring BGP on Cisco Routers* (BGP) v3.0 course

 Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Exercise Commands

Command	Description
interface type *number*	Configures an interface type and enters interface configuration mode
no shutdown	Administratively activates an interface
ip address *ip-address mask* [**secondary**]	Sets a primary or secondary IP address for an interface
encapsulation frame-relay [**ietf**]	Enables and specifies the Frame Relay encapsulation method
frame-relay map *protocol protocol-address dlci* [**broadcast**] [**ietf**] [**cisco**]	Maps between a next-hop protocol address and DLCI[1] destination address
show interfaces [**interface-name**]	Displays the statistical information specific to an interface
router ospf *process-id*	Enables OSPF[2] routing, which places you in router configuration mode
network *ip-address wildcard-mask* **area** *area-id*	Defines an interface on which OSPF runs and defines the area ID for that interface
ip ospf network {**broadcast | non-broadcast |** {**point-to-multipoint** [**non-broadcast**]}}	Configures the OSPF network type for a specified interface

[1]DLCI = data link connection identifier
[2]OSPF = Open Shortest Path First

Job Aids

These job aids are available to help you complete the laboratory exercise:

- The laboratory is organized as a number of workgroups connected to two common backbones:
  - Internet service provider (ISP) exchange point, also called the provider backbone, where two upstream service providers are located.
  - Client ISP backbone, where your customer is connected.
- Every workgroup has four routers named WGxR1, WGxR2, WGxR3, and WGxR4, where x is the number of the workgroup. There are also three shared routers called “Good,” “Cheap,” and “Client.”
- You will perform initial router configuration and prepare the routers for further exercises. During this procedure, configure passwords, serial interfaces, and IP addresses on each router in your student workgroup. You will also configure an IGP in your workgroup.

Copyright © 2003, Cisco Systems, Inc.

- Figure 1 displays the physical connectivity within your student workgroup. You have control over routers WGxR1 through WGxR4. You can also Telnet to other routers shown in the figure, but you cannot configure them.

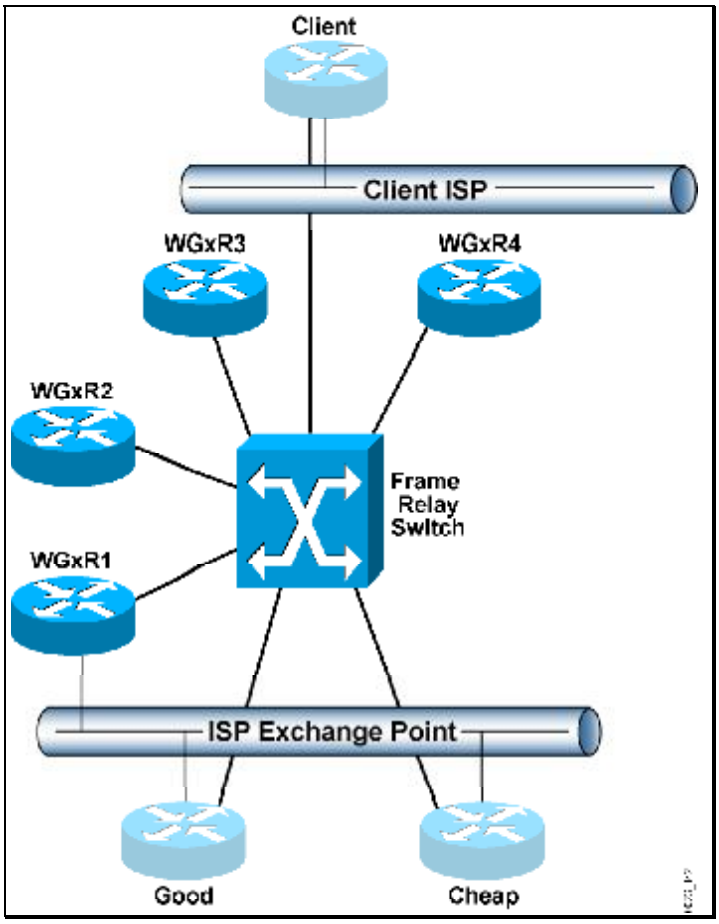


Figure 1: Physical connectivity

- The first serial interface of each of your routers is connected to the Frame Relay switch. The first (fast) Ethernet interface of each router is connected to the LAN segment. All routers including the shared ones (Good, Cheap, and Client) have one serial link to the Frame Relay switch.

- Figure 2 displays the logical connectivity of your student workgroup. Frame Relay DLCIs are already configured on the Frame Relay switch to provide this topology.

 Copyright © 2003, Cisco Systems, Inc.

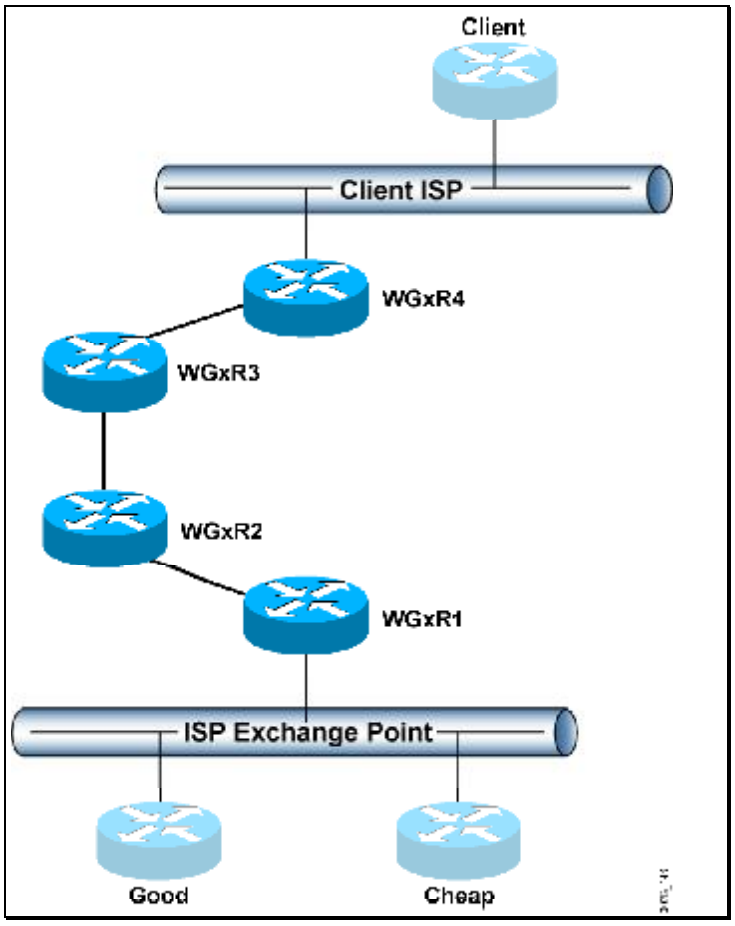


Figure 2: Logical connectivity

Exercise Procedure

Complete these steps:

Step 1 Perform initial configuration of your routers using the parameters in the following table:

Parameter	Value
host name	Use host names as shown in Figure 2 above (x is the number of your workgroup).
Enable password	Cisco
VTY password	Cisco
WAN link encapsulation	Frame Relay
WAN link clock rate	128 kbps (configured on the Frame Relay switch)

Copyright © 2003, Cisco Systems, Inc.

Step 2 Configure two loopback addresses on each of your workgroup routers with the IP addresses from the following table.

Router	Interface	Address	Subnet Mask
WGxR1	Loopback 0	197.x.1.1	255.255.255.0
	Loopback 1	197.x.8.1	255.255.255.0
WGxR2	Loopback 0	197.x.2.1	255.255.255.0
	Loopback 1	197.x.3.1	255.255.255.0
WGxR3	Loopback 0	197.x.4.1	255.255.255.0
	Loopback 1	197.x.5.1	255.255.255.0
WGxR4	Loopback 0	197.x.6.1	255.255.255.0
	Loopback 1	197.x.7.1	255.255.255.0

Step 3 Configure LAN IP addresses on WGxR1 and WGxR4 using parameters from the following table:

Parameter	Value
ISP exchange point subnet	192.168.20.x, subnet mask 255.255.255.0
Client ISP subnet	192.168.21.x, subnet mask 255.255.255.0

Note Router “Good” has IP address 192.168.20.20, and router “Cheap” has IP address 192.168.20.22. They are shared by all workgroups. Router “Client” has IP address 192.168.21.99 and is shared by all workgroups. Frame Relay DLCIs have the same value on both ends of the link.

Step 4 Configure point-to-point Frame Relay subinterfaces on the Frame Relay links. The IP addresses to be used on the link, as well as the DLCI values for the Frame Relay virtual circuits, are shown in the following table:

First router	IP address	Second router	IP address	DLCI
WGxR1	192.168.x.1/30	WGxR2	192.168.x.2/30	100
WGxR2	192.168.x.5/30	WGxR3	192.168.x.6/30	101
WGxR3	192.168.x.9/30	WGxR4	192.168.x.10/30	102

Step 5 Configure **ip host** mappings to ease Telnet hopping between routers.

Step 6 Configure any IGP between your routers. Make sure you do not use the IGP on the backbone LANs.

Note It is preferred that you use a classless IGP for this step. OSPF and Intermediate System-to-Intermediate System (IS-IS) are good choices.

 Copyright © 2003, Cisco Systems, Inc.

Exercise Verification

You have completed this exercise when you attain these results:

- All router interfaces should be active (line up, line protocol up).
- You should be able to Telnet and ping between all core routers.

Laboratory Exercise: BGP Overview

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP, given a network consisting of multiple domains.

After completing this exercise, you will be able to:

- Configure initial BGP setup
- Configure BGP neighbors
- Announce local networks in BGP
- Redistribute routes into BGP
- Configure basic BGP route aggregation

 Copyright © 2003, Cisco Systems, Inc.

- Monitor the status of the BGP routing process
- Monitor BGP neighbors
- Monitor the BGP table

Command List

The commands used in this exercise are described in the table here.

Table 1: Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor *ip-address* **remote-as** *as-number*	Establish a BGP session by using your workgroup number as the AS number.
network *network* [**mask** *mask*]	Announce IP prefixes in BGP.
ip route *network mask …*	Configure a static IP route.
route-map *name* {**permit | deny**} **seq**	Define or modify an existing entry in a route-map.
match ip address *acl*	Match routes in a route-map.
set origin igp	Set origin in a route-map.
redistribute *igp pid* **route-map** *name*	Redistribute from your IGP into BGP.
aggregate-address *network mask* [**summary-only**]	Create summary prefixes. Use the **summary-only** keyword to suppress more specific prefixes.
show ip bgp summary	Verify if the BGP session is up.
show ip bgp neighbor	View detailed information about the neighbor.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp *network*	View detailed information about prefixes (aggregates).

Job Aids

These job aids are available to help you complete the laboratory exercise:

- You must connect to the Internet using BGP, ensuring that all users in your network will get Internet access. You will connect to a single service provider and statically announce the address space that the Internet Registry has assigned to you.
- Figure 1 displays the BGP session that you have to establish between WGxR1 and the “Good” service provider.

Copyright © 2003, Cisco Systems, Inc.

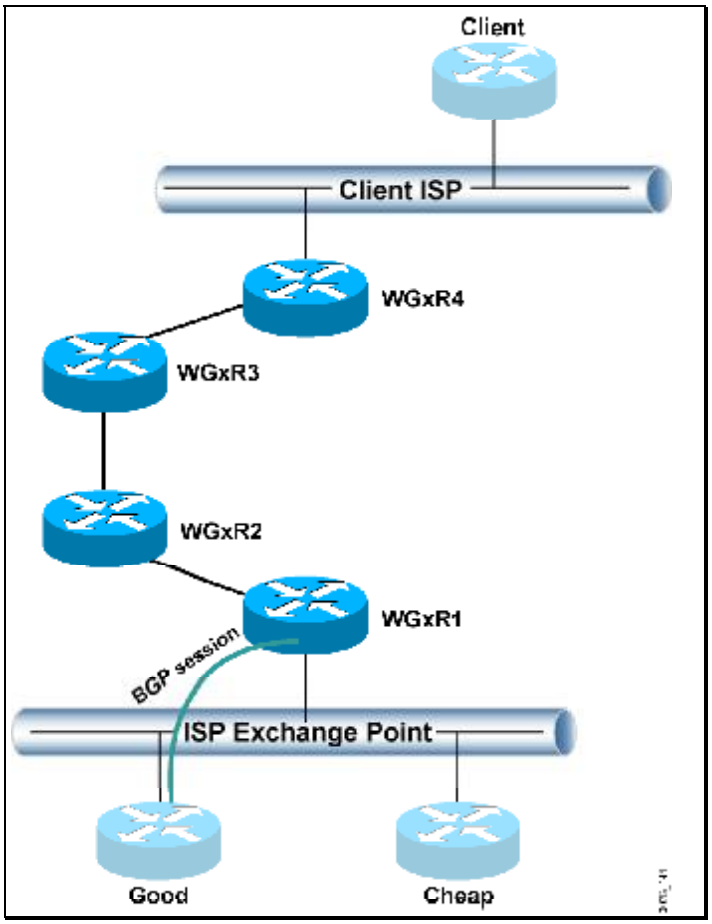


Figure 1: Connecting to single service provider

 Copyright © 2003, Cisco Systems, Inc.

Task 1: Configuring BGP

In this task, you will configure your network backbone for basic BGP connectivity with a service provider to establish BGP peering.

Exercise Procedure

Complete these steps:

Step 1 Start the BGP process on WGxR1. Use your workgroup number as the AS number.

Step 2 Configure the "Good" router to be your BGP neighbor using the following parameters.

Parameter	Value
Service provider IP address	192.168.20.20
Service provider AS number	20

Step 3 Announce the IP prefix 197.x.0.0/16 by configuring it in the BGP routing process. Also announce network 192.168.x.0/24.

Exercise Verification

You have completed this exercise when you attain these results:

- Verify that you have established a BGP session. Remember that it may take up to a minute to establish a BGP session.

```
WG1R1#show ip bgp summary
BGP router identifier 192.168.20.1, local AS number 1
BGP table version is 25, main routing table version 25
24 network entries and 24 paths using 3192 bytes of memory
5 BGP path attribute entries using 260 bytes of memory
4 BGP AS-PATH entries using 96 bytes of memory
1 BGP community entries using 250 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP activity 24/0 prefixes, 24/0 paths, scan interval 15 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxR
192.168.20.20   4    20       9       4       25    0    0 00:01:31
24
WG1R1#
```

- View detailed information about the neighbor.

```
WG1R1#show ip bgp neighbor 192.168.20.20
BGP neighbor is 192.168.20.20,  remote AS 20, external link
  BGP version 4, remote router ID 199.199.199.199
  BGP state = Established, up for 00:19:50
  Last read 00:00:50, hold time is 180, keepalive interval is 60
seconds
  Neighbor capabilities:
    Route refresh: advertised and received
    Address family IPv4 Unicast: advertised and received
  Received 27 messages, 0 notifications, 0 in queue
  Sent 22 messages, 0 notifications, 0 in queue
```

```
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 30 seconds

 For address family: IPv4 Unicast
  BGP table version 25, neighbor version 25
  Index 1, Offset 0, Mask 0x2
  24 accepted prefixes consume 864 bytes
  Prefix advertised 0, suppressed 0, withdrawn 0

  Connections established 1; dropped 0
  Last reset never
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 192.168.20.1, Local port: 179
Foreign host: 192.168.20.20, Foreign port: 18395

Enqueued packets for retransmit: 0, input: 0  mis-ordered: 0 (0
bytes)
… rest deleted …
```

- Inspect the contents of the BGP table on your router. You should see a large number of networks being advertised by the "Good" provider.

```
WG1R1#show ip bgp
BGP table version is 32, local router ID is 192.168.20.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 128.20.0.0       192.168.20.20            0             0 20 i
*> 128.22.0.0       192.168.20.20            0             0 20 i
*> 128.26.0.0       192.168.20.20            0             0 20 42 26
i
*> 128.37.0.0       192.168.20.20            0             0 20 42 37
i
*> 128.42.0.0       192.168.20.20            0             0 20 42 i
*> 128.51.0.0       192.168.20.20            0             0 20 42 26
51 i
*> 128.213.0.0      192.168.20.20            0             0 20 213 i
*> 128.214.0.0      192.168.20.22                          0 20 22
214 i
*> 192.20.11.0      192.168.20.20            0             0 20 i
*> 192.22.11.0      192.168.20.20            0             0 20 i
*> 192.26.11.0      192.168.20.20            0             0 20 42 26
i
*> 192.37.11.0      192.168.20.20            0             0 20 42 37
i
*> 192.42.11.0      192.168.20.20            0             0 20 42 i
*> 192.51.11.0      192.168.20.20            0             0 20 42 26
51 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
*> 192.168.2.0      192.168.20.2                           0 20 2 i
*> 192.168.3.2/32   192.168.20.3                           0 20 3 ?
*> 192.168.3.3/32   192.168.20.3                           0
   Network          Next Hop            Metric LocPrf Weight
*> 192.168.3.4/32   192.168.20.3                           0
*> 192.168.3.8/30   192.168.20.3                           0
*> 192.168.3.12/30  192.168.20.3                           0
*> 192.168.3.16/30  192.168.20.3                           0
*> 192.168.3.20/30  192.168.20.3                           0 20 3 ?
*> 192.213.11.0     192.168.20.20            0             0 20 213 i
*> 192.214.11.0     192.168.20.22                          0 20 2
214 i
*> 197.1.0.0/16     0.0.0.0                  0         32768 i
*> 197.2.0.0/16     192.168.20.2                           0 20 2 1
WG1R1#
```

Networks originating in AS 3. Origin code is "incomplete" because networks are redistributed.

Network originating in AS 2

 Copyright © 2003, Cisco Systems, Inc.

- Telnet from WGxR1 into the router “Good” and verify that it is receiving your networks over BGP.

- Verify that you receive networks announced by other customers.

```
*> 192.168.1.0      192.168.20.1            0             0 1 i
*> 197.1.0.0/16     192.168.20.1            0             0 1 I
```

- Perform ping and trace from WGxR4 to 192.20.11.1 (an Internet destination announced by router “Good”).

Answer these questions:

Q1) What do you need in order to propagate classful networks?

Q2) What do you need in order to propagate classless networks (supernets or subnets)?

Q3) Why do some networks, received from router “Good,” have a next-hop address pointing to other routers?

Q4) What command would you use to see if a neighbor is sending you any updates and how many?

Copyright © 2003, Cisco Systems, Inc.

Task 2: Configuring Route Redistribution in BGP

Your network has grown, and you can no longer rely on manually configuring your address space in the BGP process. In this task, you will use redistribution to announce a large number of networks into the BGP routing process. You will also use route-maps to set the origin of BGP routes to "IGP" instead of "incomplete."

In this task, you will remove all networks from your BGP definitions (from Task 1) and announce them by using redistribution from your IGP into BGP with a route-map, which sets the origin code to “IGP.” Make sure that you do not announce 192.168.20.0/24 and 192.168.21.0/24 networks into BGP.

Exercise Procedure

Complete these steps:

Step 1 Remove all BGP network statements from the previous exercise.

Step 2 Telnet to router “Good” and verify that it no longer receives your networks.

Step 3 Configure an access-list that permits all your networks except those that are shared among workgroups.

Step 4 Configure a route-map. Use the new access-list with a **match** command in the route-map. Use the **set** command in the route-map to set the origin to “IGP.”

Note	Route-maps will be covered in detail in the Route-Maps as BGP Filters lesson.

Step 5 Configure redistribution from your IGP into BGP by using the previously configured route-map.

 Copyright © 2003, Cisco Systems, Inc.

Exercise Verification

You have completed this exercise when you attain these results:

- Log into the service provider router ("Good") and verify that it receives proper networks from you.

```
Good>show ip bgp
BGP table version is 70, local router ID is 199.199.199.199
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
. . .
*> 192.168.1.0/30   192.168.20.1             0             0 1 i
*> 192.168.1.0      192.168.20.1             0             0 1 i
*> 192.168.1.4/30   192.168.20.1       2681856             0 1 i
*> 192.168.1.8/30   192.168.20.1       3193856             0 1 i
*> 197.1.1.0        192.168.20.1             0             0 1 i
*> 197.1.2.0        192.168.20.1       2297856             0 1 i
*> 197.1.3.0        192.168.20.1       2297856             0 1 i
*> 197.1.4.0        192.168.20.1       2809856             0 1 i
*> 197.1.5.0        192.168.20.1       2809856             0 1 i
*> 197.1.6.0        192.168.20.1       3321856             0 1 i
*> 197.1.7.0        192.168.20.1       3321856             0 1 i
*> 197.1.8.0        192.168.20.1             0             0 1 i
. . .
Good>
```

Individual subnets are seen if "**no auto-summary**" command is used in BGP.

- Make sure that you are not originating networks 192.168.20.0/24 and 192.168.21.0/24.

- Verify that your networks are removed from the BGP table when they become unavailable (try shutting down one of the loopback interfaces).

Answer these questions:

Q1) What is the major difference between this implementation and the previous one? Which is better and why?

Q2) What precautions do you have to take when using redistribution?

Task 3: Configuring BGP Aggregation

Your ISP has requested that you provide only summarized prefixes for your address range 197.x.0.0. However, due to diagnostic needs, you still need to announce the network 197.x.8.0.

In this task, you will configure BGP aggregation as requested, using the **aggregate-address** command.

Exercise Procedure

Complete these steps:

Step 1 Log into the router "Good"; verify that the ISP sees all your individual loopback networks.

```
> 197.1.1.0        192.168.20.1             0             0 1 i
> 197.1.2.0        192.168.20.1             0             0 1 i
> 197.1.3.0        192.168.20.1             0             0 1 i
> 197.1.4.0        192.168.20.1             0             0 1 i
> 197.1.5.0        192.168.20.1             0             0 1 i
> 197.1.6.0        192.168.20.1             0             0 1 i
> 197.1.7.0        192.168.20.1             0             0 1 i
> 197.1.8.0        192.168.20.1             0             0 1 i
```

Step 2 Use the **aggregate** command in your BGP process to aggregate the 197.x.0.0 network as requested by your ISP.

Note	Use the keyword **summary-only** if you do not want to announce individual prefixes.

Exercise Verification

You have completed this exercise when you attain these results:

- Check the BGP table on your router WGxR1 to verify the correct routing information is present.
- Log into the service providers ("Good") router and check its BGP table. Suppressed networks should not be visible on router "Good."
- View detailed information about one of the aggregates.

Answer these questions:

Q1) Do you see all your prefixes on the provider router? Why?

Q2) What do you need to be able to generate and propagate aggregate routes?

 Copyright © 2003, Cisco Systems, Inc.

2

Route Selection Using Policy Controls

Overview

The Border Gateway Protocol (BGP) enables traffic in Internet backbones to determine an optimal path to its destination across networks comprising more than one autonomous system (AS). Routes learned via BGP have properties associated with them that aid BGP in determining the best route to a particular destination. There are many instances where the default BGP route selection does not match administrative or business policies. Likewise, redundant network designs often require enterprises to run BGP when connected to more than one Internet service provider (ISP). In these situations, full BGP routing tables and default BGP route selection are not desirable.

This module provides information on how to connect Internet customers to multiple service providers. It introduces the need for filtering of BGP updates and changing of BGP route selection policies. In addition, this module describes different Cisco IOS® mechanisms (AS-path filters, prefix-lists, route-maps) available for BGP route filtering.

Upon completing this module, you will be able to:

- Describe the need for influencing BGP route selection, given a customer scenario where connections to multiple ISPs must be supported.
- Successfully configure BGP to influence route selection using AS-path filters, given a customer scenario where connections to multiple ISPs must be supported.
- Successfully configure BGP to influence route selection using prefix list filters, given a customer scenario where connections to multiple ISPs must be supported.
- Use outbound route filtering to minimize the impact of BGP routing updates on router resources, given an operational BGP network.

- Correctly configure BGP to influence route selection using route-maps, given a typical BGP network.
- Configure the soft reconfiguration feature to minimize the impact of expediting BGP policy updates, given a typical BGP network.

Outline

The module contains these lessons:

- Multihomed BGP Networks
- AS-Path Filters
- Prefix List Filters
- Outbound Route Filtering
- Route-Maps as BGP Filters
- Implementing Changes in BGP Policy

 Copyright © 2003, Cisco Systems, Inc.

Multihomed BGP Networks

Overview

Mission-critical applications often call for redundant network designs. When access to applications is provided over the Internet, enterprises typically use multihomed Border Gateway Protocol (BGP) networks to achieve their goals of high availability. In these situations, full BGP routing tables and default BGP route selection are desirable characteristics of the network designed for mission critical applications. However, the overhead of full BGP routing tables is not warranted in these situations. Further, the default route selection in BGP often does not match the business and technical requirements for multihomed enterprise networks using BGP.

This lesson discusses these business and technical issues and the requirement to use filters to influence route selection and apply policy.

Importance

In some circumstances, it is important to have multiple paths to an Internet service provider (ISP). There are business and technical reasons to configure a BGP network in a multihomed configuration. In addition to simply multihoming a BGP network, path selection is an important issue to be considered, as well as filtering particular routing advertisements.

Objectives

Upon completing this lesson, you will be able to:

- Describe the business issues of multihomed BGP networks in service provider environments
- Describe the technical issues of multihomed BGP networks in service provider environments
- Describe the need for BGP policies that influence route selection in a multihomed BGP network

- List typical routing policies for multihomed BGP customers
- Describe the need to influence BGP route selection in a service provider environment
- Describe the need for BGP filters in a service provider environment

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

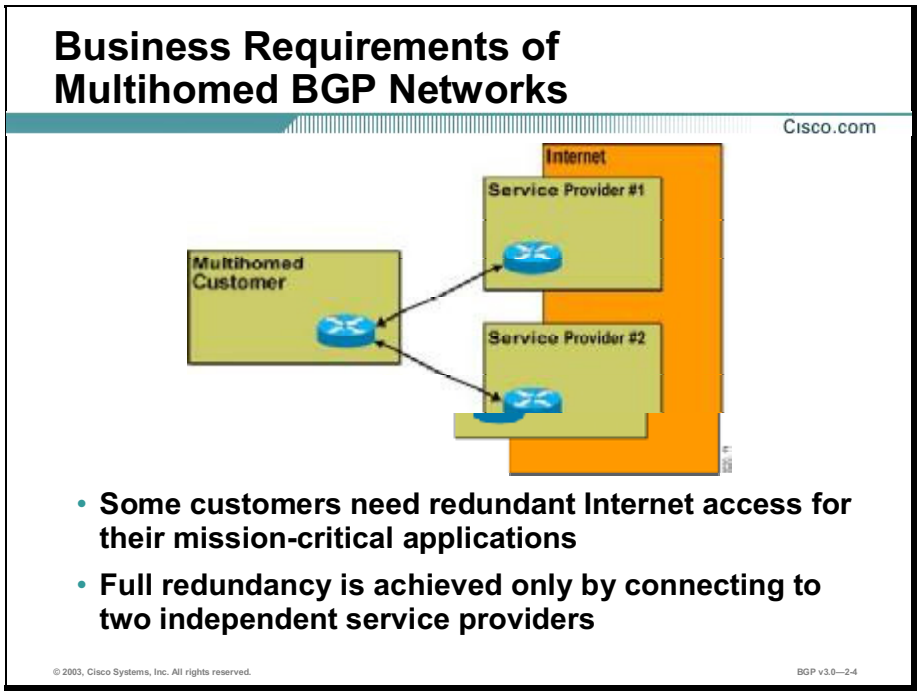
This lesson includes these topics:

- Overview
- Business Requirements of Multihomed BGP Networks
- Technical Requirements of Multihomed BGP Networks
- BGP Route Selection Without Policies
- Multihomed Customer Routing Policies
- Influencing BGP Route Selection
- BGP Filters
- Summary
- Assessment (Lab): Multihomed BGP Networks

 Copyright © 2003, Cisco Systems, Inc.

Business Requirements of Multihomed BGP Networks

This topic describes the business issues of multihomed BGP networks in service provider environments.


Business Requirements of Multihomed BGP Networks
Cisco.com
Internet
Service Provider #1
Multihomed Customer
Service Provider #2
• Some customers need redundant Internet access for their mission-critical applications
• Full redundancy is achieved only by connecting to two independent service providers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-4

Companies with web servers (or similar servers) offering mission-critical business services over the Internet often like to have their networks redundantly connected to the Internet. When the companies calculate the expected loss of business due to an unexpected disconnection from the Internet, they may conclude that having two connections to the Internet is profitable.

In such cases the company may consider being a customer to two different ISPs or having two separate connections to one ISP.

With two connections to one single ISP, BGP is usually not required. This solution provides backup for link failure and router failure. However, it does not provide backup for problems in the network of the ISP or the connection of the ISP to the rest of the Internet.

Full redundancy is achieved only by connecting to two independent ISPs. If one of the ISP networks loses its connection to the rest of the Internet, the customer will still reach the rest of the Internet via the other service provider. At the same time, the customer will still reach those users directly connected to the failing ISP via its direct connection.

Practice

Q1) What are two reasons why a customer would want to connect to two ISPs? (Choose two.)

A) to expand capacity for Internet traffic

B) to better protect confidential information as it travels through the Internet

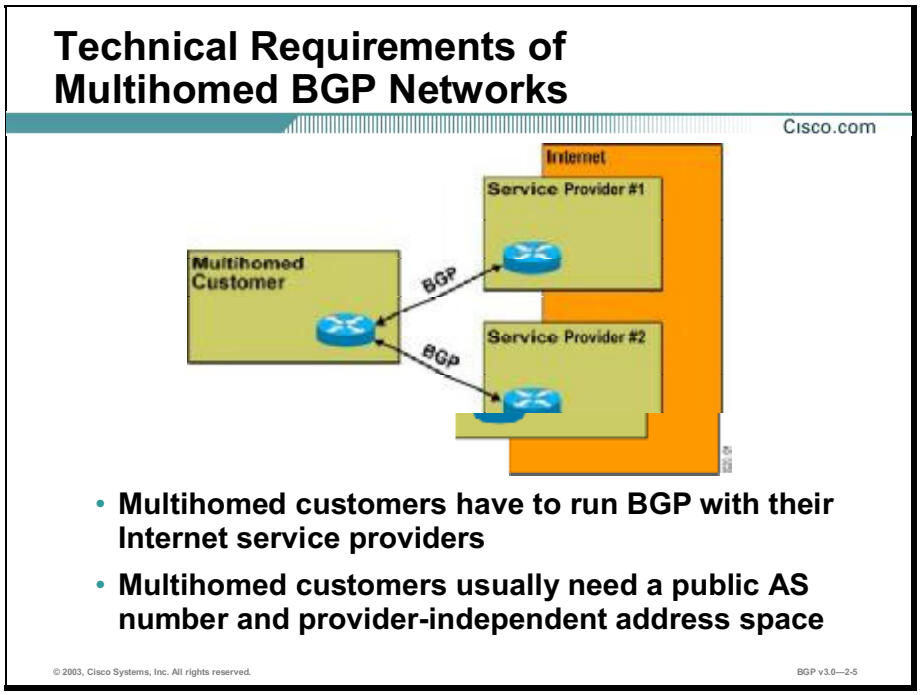
C) to provide redundancy to mission-critical services offered over the Internet

D) to efficiently route Internet traffic to two different divisions within the company

 Copyright © 2003, Cisco Systems, Inc.

Technical Requirements of Multihomed BGP Networks

This topic describes the technical issues of multihomed BGP networks in service provider environments.


Technical Requirements of Multihomed BGP Networks
Cisco.com
Internet
Service Provider #1
Multihomed Customer
BGP
BGP
Service Provider #2
• Multihomed customers have to run BGP with their Internet service providers
• Multihomed customers usually need a public AS number and provider-independent address space
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-5

The multihomed customer network must exchange BGP information with both ISP networks. Dynamic routing is required for full redundancy, and BGP is the only protocol available that can be used in this scenario.

The customer must, in most cases, have its own public AS number and announce its own IP networks to both ISPs. The ISPs will propagate customer announcements to the rest of the Internet, and the customer will be seen as reachable via both ISP networks. The customer network also receives full Internet routing from both ISPs. This capability gives the customer network the opportunity to choose the best connection at that time to reach any destination on the Internet.

Most customers are not multihomed. They do not exchange BGP with their ISP. Instead, they use default routing to the ISP, and the ISP does static routing to the customer. ISPs use this fact to optimize the number of prefixes that they announce into the Internet. IP network numbers are usually assigned to customers from the range of IP networks delegated to the ISP to which the customer is connected. This situation means, in the ideal case, that all customers connected to one single ISP can have their IP networks summarized in a few BGP updates.

In the multihomed scenario, however, the ISP cannot benefit from IP network number assignment from the delegated range. The customer is connected to two different ISPs, and it is not obvious from which provider-assigned address space it should get the IP addresses. The best solution is to do the assignment from a range completely independent of the providers, a provider-independent address space.

Practice

Q1) What are the two technical requirements for multihomed customers? (Choose two.)

A) The ISPs must assign a range of IP network numbers to the customer.

B) The customer network must exchange BGP information with each ISP network.

C) In most cases, the customer must have its own public AS number.

D) The customer network must not use AS-path filters.

Q2) Which routing protocol needs to be deployed between multihomed customers and the ISP?

A) any interior gateway protocol

B) static routes

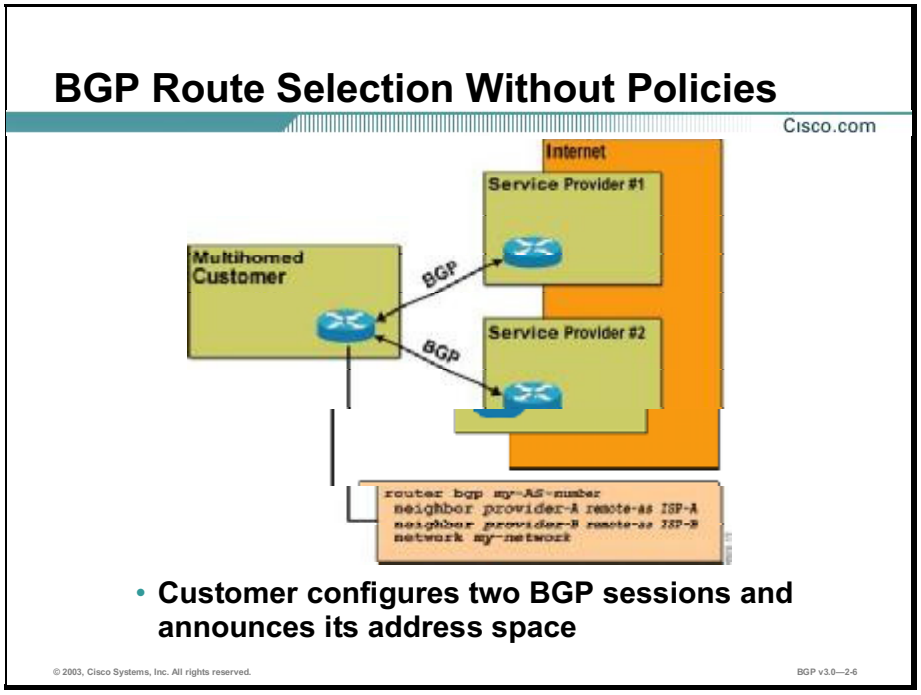
C) default routes

D) BGP

 Copyright © 2003, Cisco Systems, Inc.

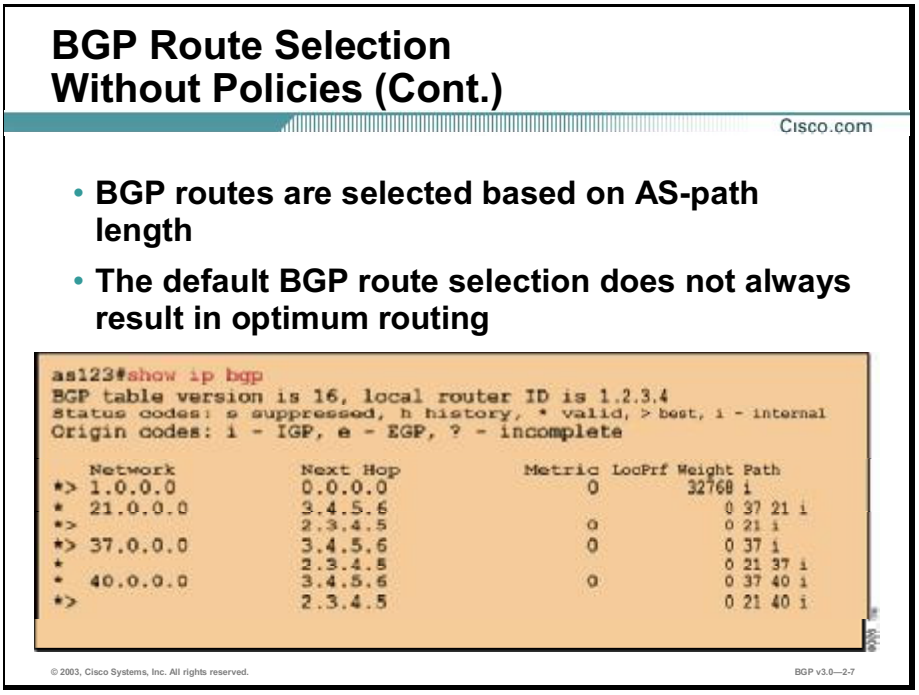
BGP Route Selection Without Policies

This topic describes the need for BGP policies that influence route selection in a multihomed BGP network.



The simple approach illustrated in the figure may be the source of many problems. By simply starting BGP sessions with both ISPs, and announcing its networks to them both, a customer could experience difficulties as a result of the default behaviors of BGP.

Example


BGP Route Selection
Without Policies (Cont.)
Cisco.com
• BGP routes are selected based on AS-path length
• The default BGP route selection does not always result in optimum routing
as123#show ip bgp
BGP table version is 16, local router ID is 1.2.3.4
Status codes: s suppressed, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 1.0.0.0 0.0.0.0 0 32768 i
* 21.0.0.0 3.4.5.6 0 37 21 i
*> 2.3.4.5 0 0 21 i
*> 37.0.0.0 3.4.5.6 0 0 37 i
* 2.3.4.5 0 21 37 i
* 40.0.0.0 3.4.5.6 0 0 37 40 i
*> 2.3.4.5 0 21 40 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-7

This is an example of a multihomed customer with AS 123. It is connected to two different ISPs: AS 37 and AS 21. The two ISPs are interconnected, and both are also connected to AS 40.

The customer receives all routes from both service providers, giving redundancy. The default BGP route selection prefers the shortest AS path. If the AS-path lengths are equal, it prefers the most stable route, or the route received from the peer with the lower router-ID.

In many cases, however, this is not the most optimal way to reach all destinations. For example, the bandwidth available to reach the ISPs has not been taken into consideration. To change the route selection behavior, some BGP parameters must be configured to support more complex routing policies.

Practice

Q1) In the absence of applied policy, what two factors will most likely influence BGP route selection? (Choose two.)

A) weight

B) AS path

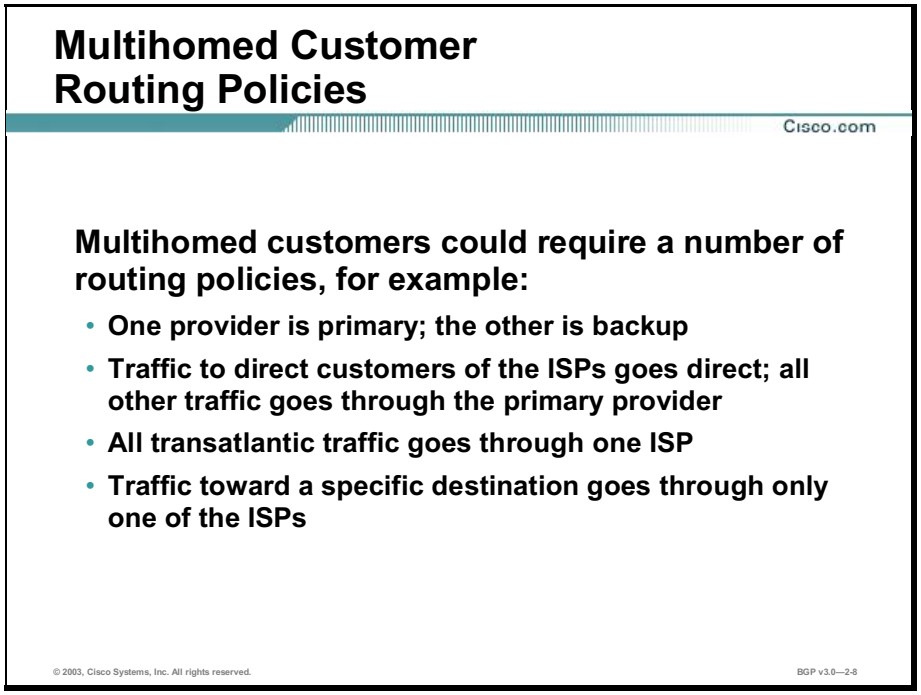
C) origin

D) router-ID

 Copyright © 2003, Cisco Systems, Inc.

Multihomed Customer Routing Policies

This topic lists typical routing policies for multihomed BGP customers.


Multihomed Customer Routing Policies
Cisco.com
Multihomed customers could require a number of routing policies, for example:
• One provider is primary; the other is backup
• Traffic to direct customers of the ISPs goes direct; all other traffic goes through the primary provider
• All transatlantic traffic goes through one ISP
• Traffic toward a specific destination goes through only one of the ISPs
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-8

Depending on the circumstances, there are a number of different polices that a multihomed customer might require:

- One of the two ISPs may be considered the primary connection. This distinction can be due to the available bandwidth or to commercial agreements. However, although one of the ISPs is considered the primary connection, some users may have direct connections to the secondary ISP. Therefore, going via the primary ISP to reach users connected to the secondary ISP may be suboptimal.

- Destinations on the other side of the Atlantic may be reached more optimally via one of the ISPs, rather than via the other, as the two ISPs may have different infrastructures and peering agreements with other ISPs.

To establish a routing policy that gives optimal routing to each and every destination on the Internet is virtually impossible. Optimization can be done only with the most common destinations in mind. This situation can result in specific rules on how to reach specific destination networks or the AS.

Practice

Q1) List three potential customer routing policies. (Choose three.)

A) One service provider is designated as primary, and the other is a backup.

B) Traffic is load-balanced across both ISP networks.

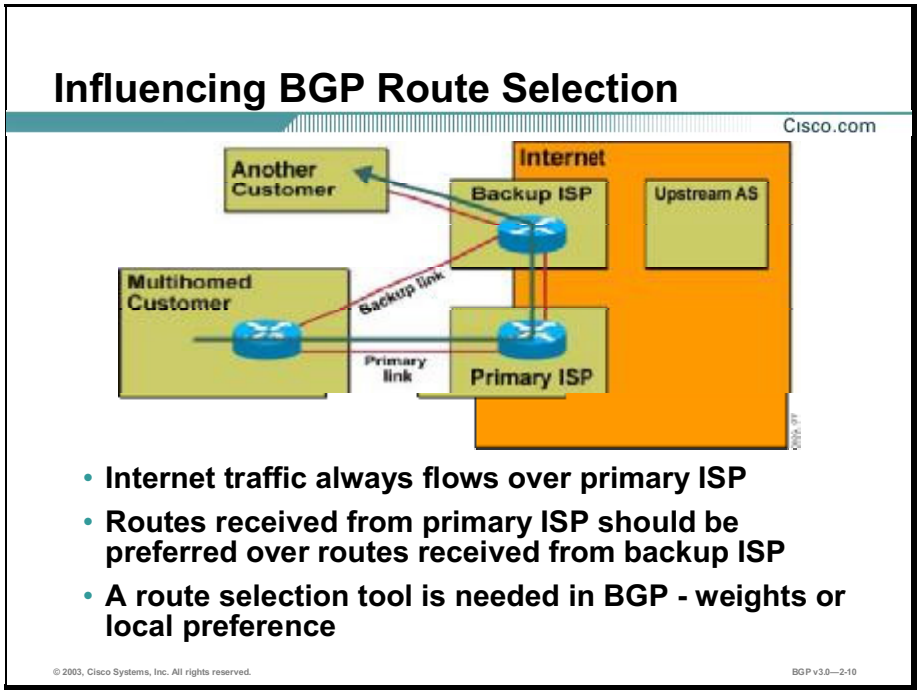
C) Traffic toward a specific destination goes through only one of the ISPs.

D) Traffic to direct customers of the ISPs goes direct; all other traffic goes through the primary ISP.

 Copyright © 2003, Cisco Systems, Inc.

Influencing BGP Route Selection

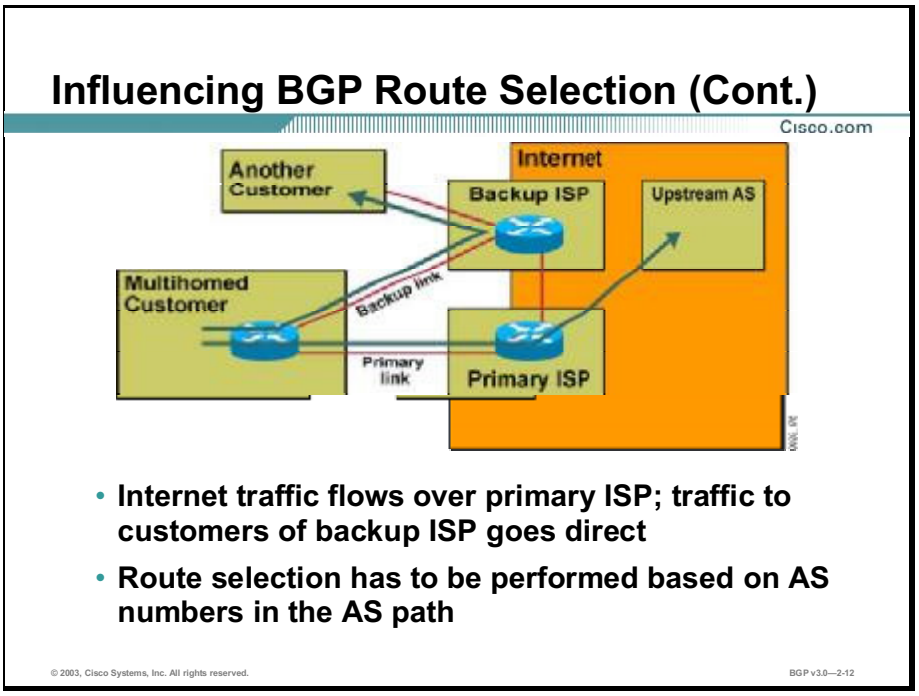
This topic describes the need to influence BGP route selection in a service provider environment.


Influencing BGP Route Selection
Cisco.com
Another Customer
Internet
Backup ISP
Upstream AS
Multihomed Customer
Backup link
Primary link
Primary ISP
• Internet traffic always flows over primary ISP
• Routes received from primary ISP should be preferred over routes received from backup ISP
• A route selection tool is needed in BGP - weights or local preference
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-10

When one of the two ISPs is designated as a primary ISP and the other a backup, BGP attributes must be configured as a means of influencing BGP route selection rules. If both ISP connections terminate in one single customer router, all routes received from the primary ISP can be assigned a BGP weight. A higher weight indicates a more preferred path.

But the weight value is local to one router. It is not shared between routers. If one ISP connection terminates in one of the customer routers, and the other ISP connection terminates in another, the two customer routers must agree on which link to use. Using local preference instead of weight can do this. All routes received from the primary ISP over the primary link are assigned a local preference value, which is higher than the default value 100. The customer router receiving the routes from the primary ISP completes the assignment and communicates the information to the other routers within the AS of the customer.

The result of using either weight or local preference is that the AS of the customer reaches all its destinations on the Internet via the primary link as long as it is available, as well as those destinations within the AS of the secondary ISP. In the case of link failure, or failures within the network of the primary ISP, some of the routes, or all of the routes, will no longer be received over the primary link. In that case, the AS of the customer no longer sees those destinations as reachable over the primary link. The only remaining choice is via the backup link. Therefore, the backup link is used by the customer network only to reach destinations that are not reachable over the primary link.


Influencing BGP Route Selection (Cont.)
Cisco.com
Another Customer
Internet
Backup ISP
Upstream AS
Multihomed Customer
Backup link
Primary link
Primary ISP
• Internet traffic flows over primary ISP; traffic to customers of backup ISP goes direct
• Route selection has to be performed based on AS numbers in the AS path
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-12

In most cases, it is more optimal to reach other customers connected to the backup ISP via the backup link, compared with reaching them via the primary link.

The routing policy, described previously, where routes are blindly preferred if received on the primary link, can easily be modified to use the backup link when reaching destinations in the AS of the backup ISP. On the router, filtering tools can be configured to select routing information based on the content in the AS-path attribute. Those routes, with an AS-path matching specific selection criteria, can be assigned an even higher weight or local preference.

This approach results in a routing policy that gives precedence to reaching destinations within the AS of the primary ISP and within all autonomous systems upstream of the primary ISP over the primary link. Destinations within the AS of the backup ISP receive precedence over the backup link.

Practice

Q1) Why do you need to influence BGP route selection rules?

A) because the default BGP route selection does not always result in optimum routing

B) because BGP route selection favors the path with the highest bandwidth, which is not always the best path

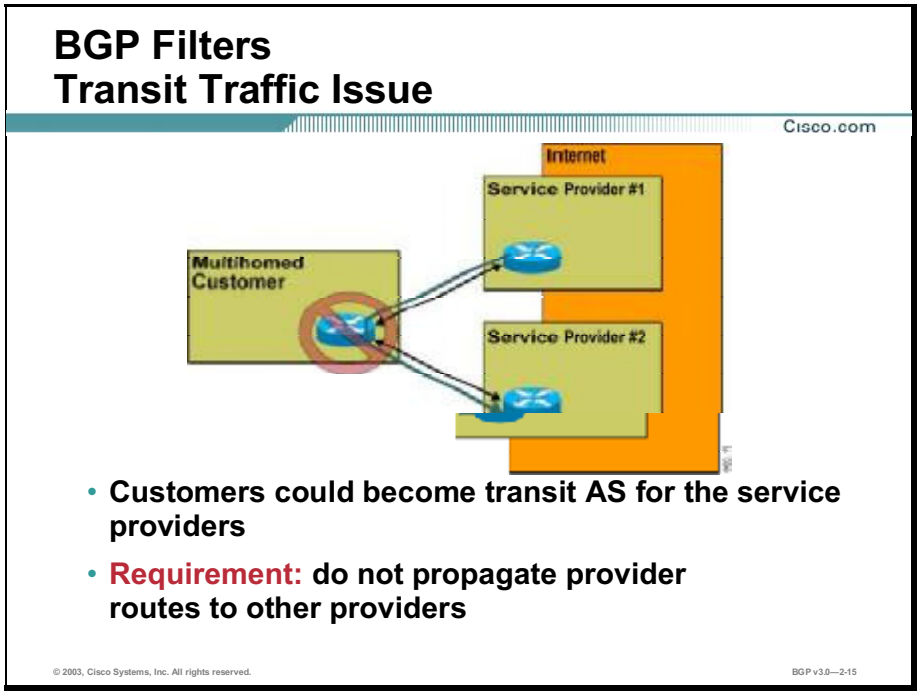
C) because the ISP will select the route if you do not influence BGP route selection

D) because BGP route selection often favors less stable routes

 Copyright © 2003, Cisco Systems, Inc.

BGP Filters

This topic describes the need for BGP filters in a service provider environment.

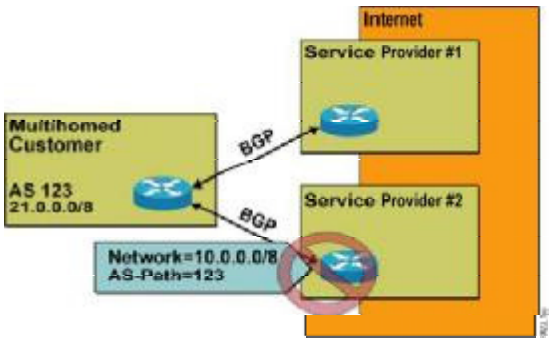

BGP Filters
Transit Traffic Issue
Cisco.com
Internet
Service Provider #1
Multihomed Customer
Service Provider #2
• Customers could become transit AS for the service providers
• Requirement: do not propagate provider routes to other providers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-15

When BGP has selected the best path, the information is advertised by the router to all neighboring autonomous systems, except on the session it was received on (called BGP split-horizon functionality, preventing near-range routing loops). This causes the customer AS to become a transit between the two ISPs. This situation should be avoided.

Most customers do not have the intention to transit traffic between ISP networks. The access lines to the ISPs are not suited to carry this volume of traffic, and the customer certainly does not want to have its bandwidth consumed by transit traffic.

The solution to this problem is to filter outgoing information to both ISPs. Filtering of routing information is performed based on the content of the AS-path attribute assigned to every BGP route. Only those routes having an AS-path attribute indicating that they are sourced by the AS of the customer are allowed to be sent to either of the two ISPs.

BGP Filters
Routing Update Reliability Issue



- **Customers running BGP could announce any route to the service providers**
- **Requirement: service providers have to filter IP prefixes in incoming updates**

© 2003, Cisco Systems, Inc. All rights reserved.

Without some sort of filtering, BGP routing information created by the AS of the customer can potentially be propagated all over the Internet. In this way, the customer can inject erroneous information into the Internet routing tables.

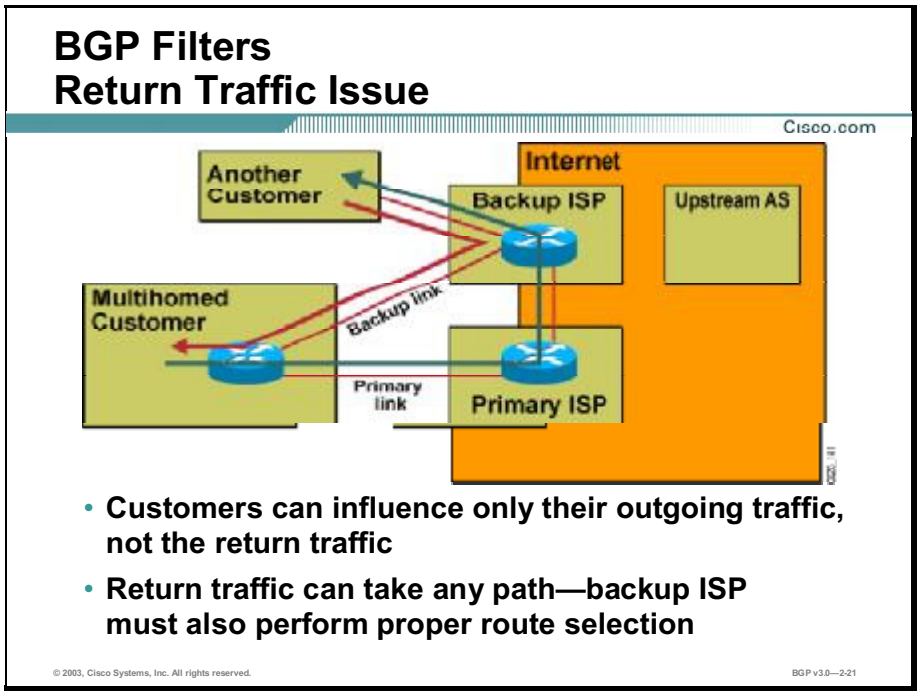
Customers are much less experienced in avoiding these kinds of problems than are service providers. There is much more risk for errors to be introduced when a customer is assigned its own AS and uses BGP with the ISP, compared with the single-homed scenario where the ISP has sole responsibility to announce BGP routes to the rest of the Internet.

Almost all the problems a customer can cause the Internet by improperly configuring its BGP can be stopped by the ISP. The ISP should filter all incoming information from the customer and accept only what is supposed to arrive. It should discard anything outside strict limits. In this way, the ISP prevents the propagation of erroneous information to the rest of the Internet.

The ISP can maintain a list of the IP network numbers that the customer is announcing and filter out any other route. If this approach is not possible due to the volume of those lists, the ISP should at least be able to filter out the most obvious erroneous announcements.

Note Private addresses, according to RFC 1918, should never be announced to the Internet.

 Copyright © 2003, Cisco Systems, Inc.



The customer can easily define a policy about how to send outgoing IP packets on the correct link. It is much harder to influence the neighboring AS about how to direct the IP packets coming into the customer network.

A customer who creates a routing policy where one of the two ISPs is always the preferred may see that the return traffic is arriving on what the customer thought was the backup link. This situation means that the customer has configured the weight or local preference to make sure that all outgoing traffic is leaving the customer AS over the primary link, but the backup ISP does not have any such configuration. Therefore, return traffic enters the customer's AS using the shortest AS-path as selection criterion.

The best way to solve this problem is for the customer to ask the backup ISP to change its routing policy. The change should cause the backup ISP to prefer reaching the customer AS via the AS of the primary ISP. The backup ISP must implement this change in its own AS.

Note	Sometimes the backup ISP administrator could be reluctant to change the configuration for a single customer. In this case, the customer should use another BGP feature, the AS-path prepending tool, to influence the selection of the primary/backup link by lengthening the AS path of routes sent to the backup provider.

Practice

Q1) What two potential multihomed network issues can be prevented with IP prefix filters? (Choose two.)

A) the propagation of private AS numbers

B) the propagation of private addresses used in the network

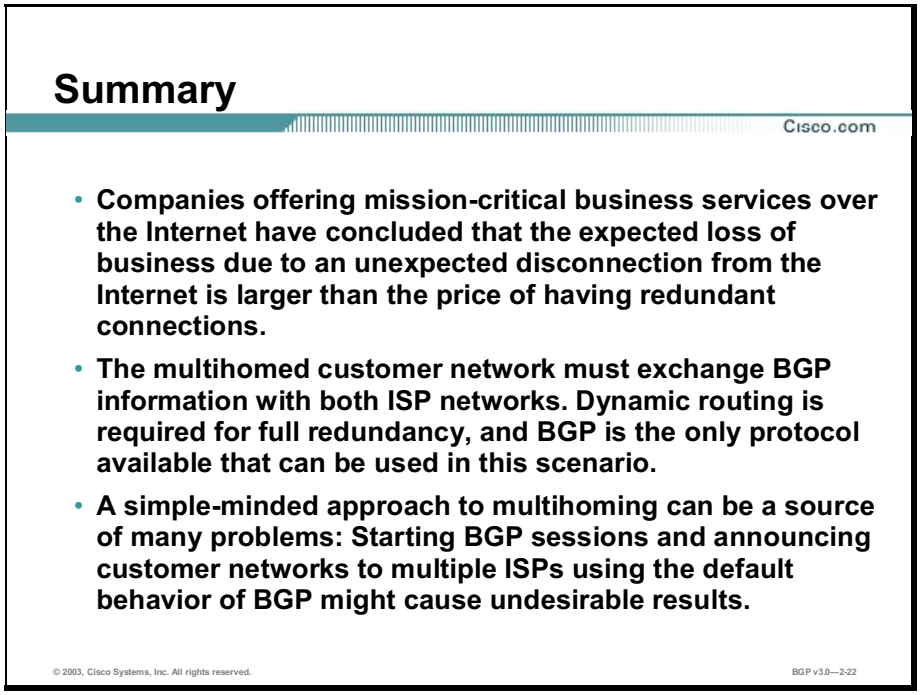
C) the propagation of unreachable next-hop addresses

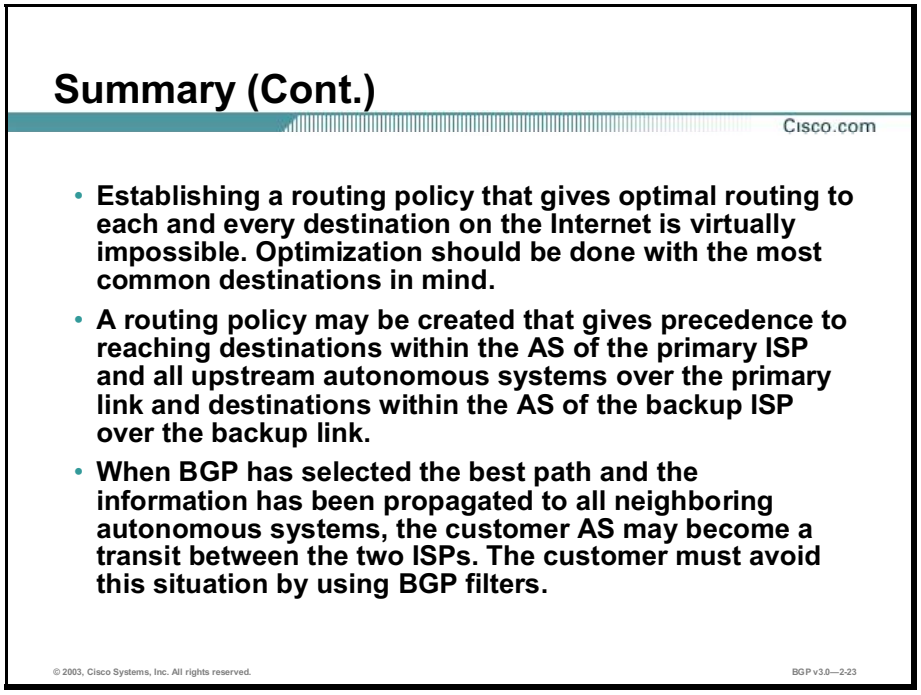
D) the propagation of more specific prefixes from an address range

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Companies offering mission-critical business services over the Internet have concluded that the expected loss of business due to an unexpected disconnection from the Internet is larger than the price of having redundant connections.
• The multihomed customer network must exchange BGP information with both ISP networks. Dynamic routing is required for full redundancy, and BGP is the only protocol available that can be used in this scenario.
• A simple-minded approach to multihoming can be a source of many problems: Starting BGP sessions and announcing customer networks to multiple ISPs using the default behavior of BGP might cause undesirable results.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-22


Summary (Cont.)
Cisco.com
• Establishing a routing policy that gives optimal routing to each and every destination on the Internet is virtually impossible. Optimization should be done with the most common destinations in mind.
• A routing policy may be created that gives precedence to reaching destinations within the AS of the primary ISP and all upstream autonomous systems over the primary link and destinations within the AS of the backup ISP over the backup link.
• When BGP has selected the best path and the information has been propagated to all neighboring autonomous systems, the customer AS may become a transit between the two ISPs. The customer must avoid this situation by using BGP filters.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-23

Next Steps

After completing this lesson, go to:

- AS-Path Filters lesson

References

For additional information, refer to these resources:

- For more information on multihomed networks using BGP, refer to “Sample Configuration for BGP with Two Different Service Providers (Multihoming)” at the following URL: http://www.cisco.com/warp/public/459/27.html

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: Multihomed BGP Networks

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP given a customer scenario where you must support connections to multiple ISPs.

After completing this exercise, you will be able to:

- Configure BGP neighbors to support a multihomed customer scenario
- Monitor the status of the BGP routing process
- Monitor BGP neighbors in a multihomed customer scenario

Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode
neighbor *ip-address* **remote-as as**	Start BGP session with the neighboring AS
neighbor ip-address weight *weight*	Assign weight to all updates received from the specified neighbor
show ip bgp summary	Verify the state of BGP sessions
show ip bgp	Inspect the contents of the BGP table

Job Aids

These job aids are available to help you complete the laboratory exercise:

- You have started to provide mission-critical e-commerce services, and you must ensure their high availability. You decide to connect to a new ISP, “Cheap,” using “Cheap” as your primary ISP and “Good” as your backup ISP.

- Figure 1 shows the connectivity that you need to establish for the second BGP session with the new provider router (“Cheap”).

 Copyright © 2003, Cisco Systems, Inc.

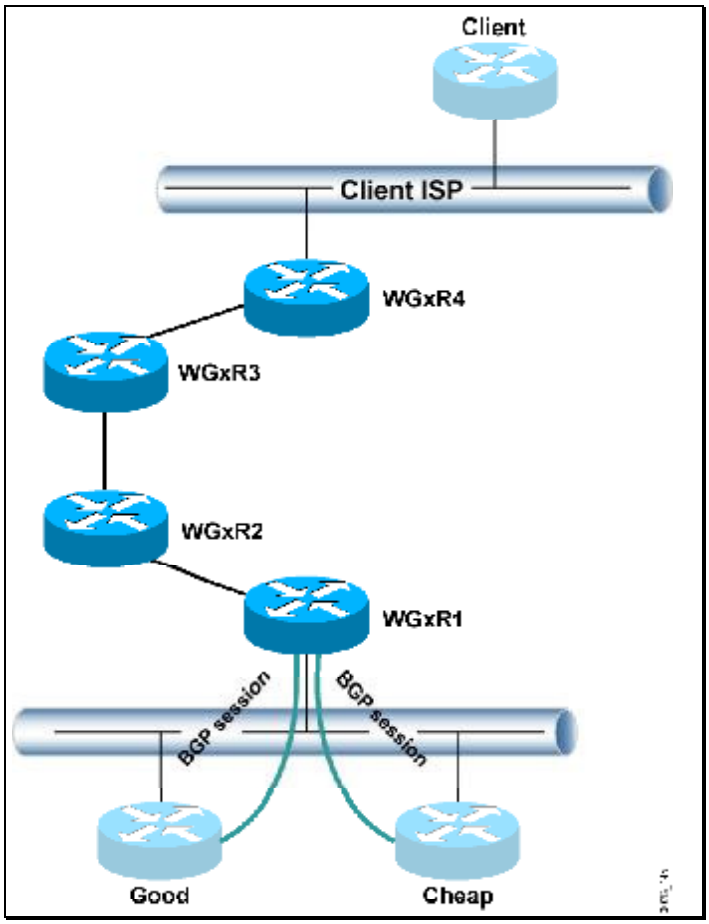


Figure 1: Connecting to two different service providers

Exercise Procedure

Complete these steps:

Step 1 Configure the second BGP neighbor on WGxR1 using the following parameters:

Service provider	Parameter	Value
Cheap	AS number	22
	IP address	192.168.20.22

Step 2 Change the default weight on neighbor "Cheap" to 100 using the **neighbor** *ip-address* **weight** *weight* command to give preference to routes announced from router "Cheap" over those announced from router "Good."

Note BGP weights are covered in detail in the Route Selection Using Attributes module and its supporting lab exercises.

Copyright © 2003, Cisco Systems, Inc.

Exercise Verification

You have completed this exercise when you attain these results:

- Verify that both BGP sessions are up.

```
WG1R1#sh ip bgp summary
BGP router identifier 197.1.8.1, local AS number 1
BGP table version is 122, main routing table version 122
35 network entries and 58 paths using 5483 bytes of memory
18 BGP path attribute entries using 936 bytes o
16 BGP AS-PATH entries using 384 bytes of memor
0 BGP community entries using 0 bytes of memory
0 BGP route-map cache entries using 0 bytes of
0 BGP filter-list cache entries using 0 bytes of memory
BGP activity 49/464 prefixes, 95/37 paths, scan interval 15 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxR
192.168.20.20   4    20     411     412      122    0    0 00:07:10
23
192.168.20.22   4    22      45      34      122    0    0 00:07:11
23
WG1R1#
```

Both sessions are up, and 23 prefixes were received from both neighbors.

- Check all the BGP prefixes with two paths to verify that your router prefers the one through router "Cheap."

```
wg1r1#sh ip bgp
BGP table version is 122, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.0.0.0         192.168.20.233                       100 22 20 i
*                   192.168.20.233           0             0 20 i
*> 128.20.0.0       192.168.20.22            0           100 22 i
*                   192.168.20.20            0             0 20 i
*> 128.20.12.0/24   192.168.20.20                        100 22 20 i
*                   192.168.20.20            0             0 20 i
*> 128.22.0.0       192.168.20.22            0           100 22 i
*                   192.168.20.20            0             0 20 i
*> 128.22.12.0/24   192.168.20.22            0           100 22 i
*                   192.168.20.22                         0 20 22 i
*> 128.26.0.0       192.168.20.22            0           100 22 26 i
*                   192.168.20.20            0             0 20 42 26
i
*> 128.37.0.0       192.168.20.22            0           100 22 26 42
37 i
*                   192.168.20.20            0             0 20 42 37
i
*> 128.42.0.0       192.168.20.22            0           100 22 26 42
i
*                   192.168.20.20            0             0 20 42 i
*> 128.51.0.0       192.168.20.22            0           100 22 26 51
i
*                   192.168.20.20            0             0 20 42 26
51 i
*> 128.213.0.0      192.168.20.20                        100 22 20
213 i
*                   192.168.20.20            0             0 20 213 i
*> 128.214.0.0      192.168.20.22            0           100 22 214 i
```

Both paths have the same next-hop address.

Copyright © 2003, Cisco Systems, Inc.

```
*                   192.168.20.22                          0 20 22
214 i
*> 192.20.11.0      192.168.20.22            0           100 22 i
*                   192.168.20.20            0             0 20 i
*> 192.20.12.0/30   192.168.20.20                        100 22 20 i
*                   192.168.20.20            0             0 20 i
*> 192.22.11.0      192.168.20.22            0           100 22 i
*                   192.168.20.20            0             0 20 i
*> 192.22.12.0/30   192.168.20.22            0           100 22 i
*                   192.168.20.22                          0 20 22 i
*> 192.26.11.0      192.168.20.22            0           100 22 26 i
*                   192.168.20.20            0             0 20 42 26
i
*> 192.37.11.0      192.168.20.22            0           100 22 26 42
37 i
*                   192.168.20.20            0             0 20 42 37
i
*> 192.42.11.0      192.168.20.22            0           100 22 26 42
i
*                   192.168.20.20            0             0 20 42 i
*> 192.51.11.0      192.168.20.22            0           100 22 26 51
i
*                   192.168.20.20            0             0 20 42 26
51 i
```

Answer these questions:

Q1) What can happen if a multihomed AS is passing routing information, learned from one neighbor, to another neighbor?

Q2) Why do some prefixes have two paths but both use the same next-hop address?

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Filters

Overview

The Border Gateway Protocol (BGP) allows connectivity between multiple Internet service providers (ISPs) for redundancy and scalability. AS-path filters are employed by service providers to remedy the problems associated with the various connectivity methods used within BGP. This lesson explains the methods used to implement BGP AS-path filters.

Importance

In network implementations requiring connections to multiple ISPs, network operators typically use AS-path filters to influence BGP route selection. It is important for a network administrator to understand the syntax of an AS-path regular expression and how string matching operators function when using AS-path regular expressions to match BGP routes.

Objectives

Upon completing this lesson, you will be able to:

- Identify network scenarios where you must support connections to multiple ISPs and AS-path filters would be used to influence route selection
- Describe the purpose of an AS-path regular expression
- Describe how string-matching operators function when you are using AS-path regular expressions to match BGP routes
- Identify where you can apply an AS-path filter when configuring a router to influence route selection
- Identify the Cisco IOS® commands required to configure AS-path filters to influence route selection

- Identify the Cisco IOS commands required to monitor the operation of configured AS-path filters

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

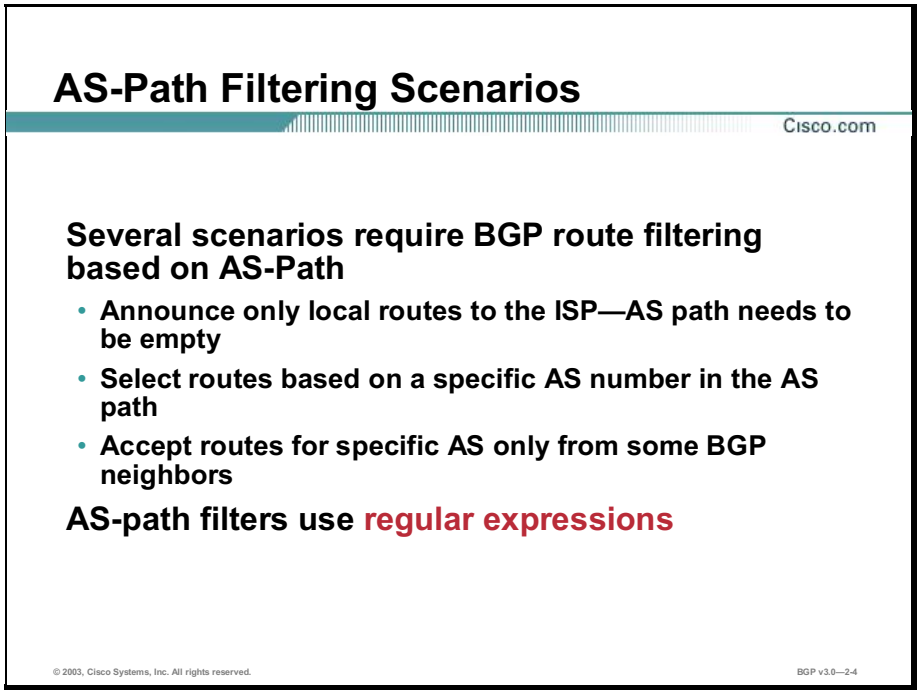
This lesson includes these topics:

- Overview
- AS-Path Filtering Scenarios
- AS-Path Regular Expressions
- String Matching
- Applying AS-Path Filters
- Configuring BGP AS-Path Filters
- Monitoring AS-Path Filters
- Summary
- Assessment (Lab): AS-Path Filters

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Filtering Scenarios

This topic identifies network scenarios that require connections to multiple ISPs where route selection must be influenced with AS-path filters.


AS-Path Filtering Scenarios
Cisco.com
Several scenarios require BGP route filtering based on AS-Path
• Announce only local routes to the ISP—AS path needs to be empty
• Select routes based on a specific AS number in the AS path
• Accept routes for specific AS only from some BGP neighbors
AS-path filters use regular expressions
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-4

Several scenarios require filtering and selection of routing information, based on the content of the AS-path attribute. Each BGP route must have an AS-path attribute. It is a well-known mandatory attribute and must therefore be present in each BGP update.

Using selection criteria based on the AS-path attribute, a router can identify a set of specific routes from the total set of routes it receives. Those routes where the AS-path contents match the criteria are selected. Routes that do not match the criteria are not selected.

The AS path is a sequence of numbers. Each number indicates an AS. When a route is sourced by means of a **network** command in a BGP process or redistribution into a BGP process, the AS-path attribute is created and left empty. Each time the route is advertised by an egress router to another AS, the AS-path attribute is modified by the egress router, which prepends its AS number to the AS-path attribute.

While a newly sourced route is still within the AS in which it was created, the AS path is still empty. So, when an AS has a requirement to filter out all but the routes local to itself before sending them to a neighboring AS, it will permit sending of the routes with the empty AS path and will deny all others.

Routers can also filter incoming routes based on their AS-path attributes. Some destination autonomous systems should not be received from a certain neighbor. Therefore, those routes matching that AS in the AS path can be filtered on the receiving router in case they are accidentally sent.

Selection based on the AS path is also a tool that can be used when changing the weight or local preference attributes for some destination autonomous systems but not for others.

When routers filtering BGP updates based on the content of the AS-path attribute, regular expressions are used. Regular expressions are commonly found in the UNIX environment and also in some Windows-based applications. They are a string-matching tool. A regular expression consists of a string of characters. Some of them have special meanings, such as to function as wildcards and operators, and some of them mean themselves, for example, A-Z, a-z, or 0-9. A regular expression is said to match a string if the ordinary characters and the applied meaning of the special operator characters can be translated into the matched string. When a regular expression matches, the selection test is said to be true. If it does not match, the test is false.

Practice

Q1) Which three of these situations are appropriate applications of AS-path filters? (Choose three.)

A) to ensure only locally originated routes are announced

B) to limit routes advertised from IBGP neighbors

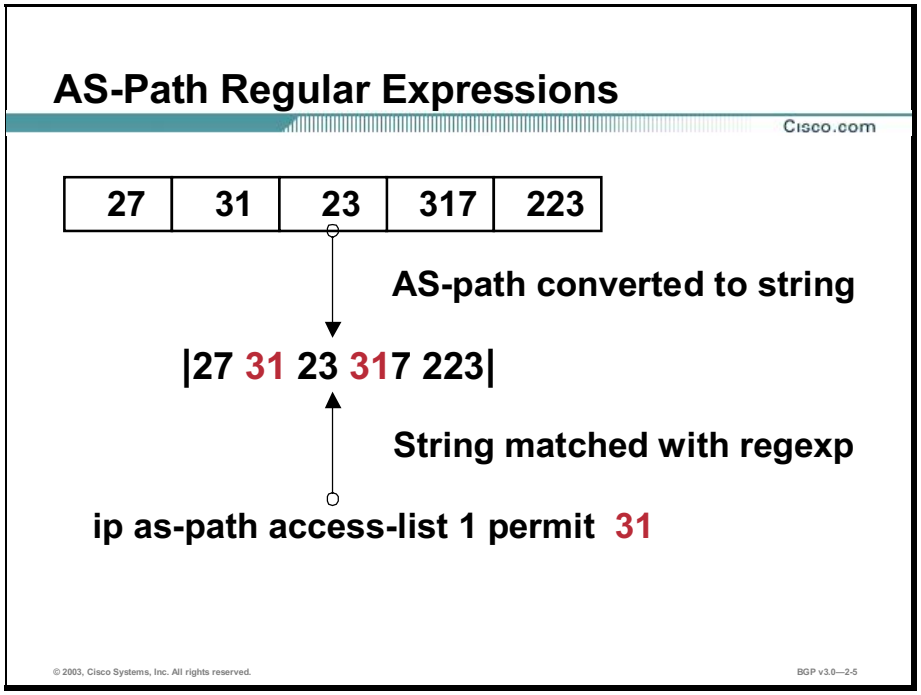
C) to select a subset of all routes based on their originating AS

D) to limit neighbor route updates to specific AS-originated routes

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Regular Expressions

This topic describes the function of an AS-path regular expression.



The AS-path attribute carried with all BGP routes in a BGP update is a very compact binary encoding of a sequence of integer numbers. It is not a string that can be tested using a regular expression.

Cisco IOS software internally translates the binary encoding into a character string. Each AS number in the sequence is converted into a string using decimal representation. The space character separates each AS number in the AS-path attribute. The router applies the regular expression test to this internally created character string.

Characters in a regular expression that are not assigned a specific operation match themselves. The regular expression “31” matches all occurrences of the character “3” followed by the character “1” in the AS path. In this example, “31” matches at two occurrences. One occurrence is sufficient to make the test true. No occurrence means that the test failed.

Practice

Q1) Which AS path is not matched by the regular expression “27”?

A) 100 27

B) 27 64 128

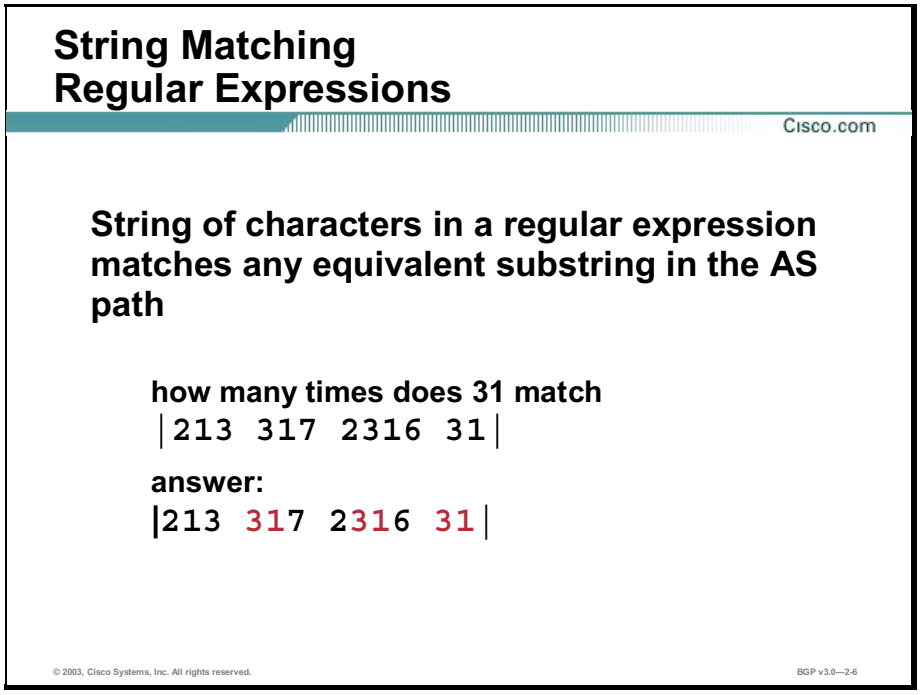
C) 10 12 182 77

D) 10 12 182 77 27 71

 Copyright © 2003, Cisco Systems, Inc.

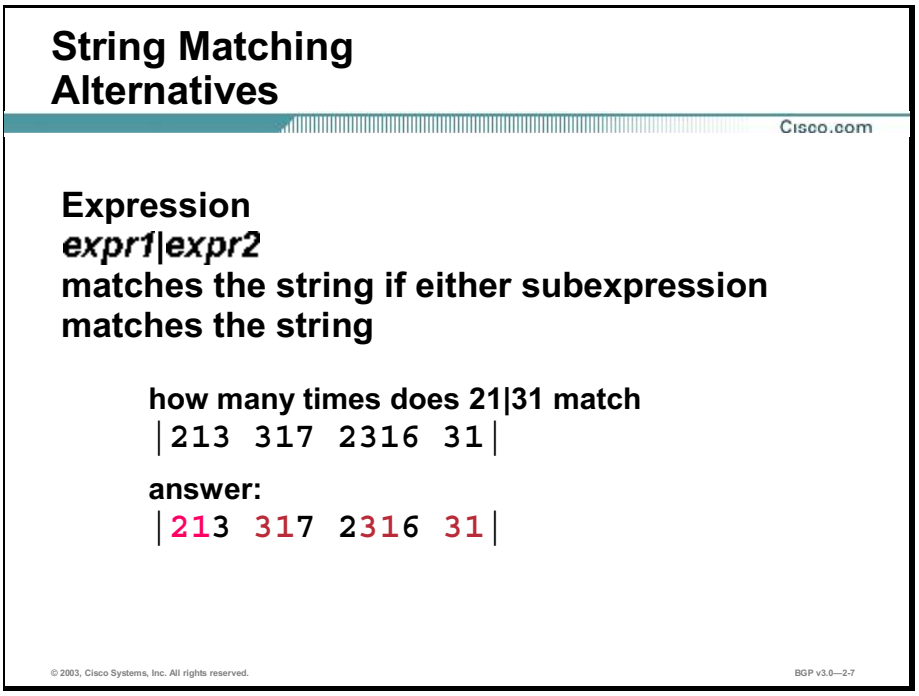
String Matching

This topic describes how string matching functions when you are using AS-path regular expressions to match BGP routes.


String Matching
Regular Expressions
Cisco.com
String of characters in a regular expression matches any equivalent substring in the AS path
how many times does 31 match
|213 317 2316 31|
answer:
|213 317 2316 31|
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-6

The regular expression “31” will match any occurrence of “3” followed by “1”, regardless of the characters immediately preceding the “3” and immediately following the “1”. So “31” will match an occurrence of “3” and “1” in the middle of an AS number.

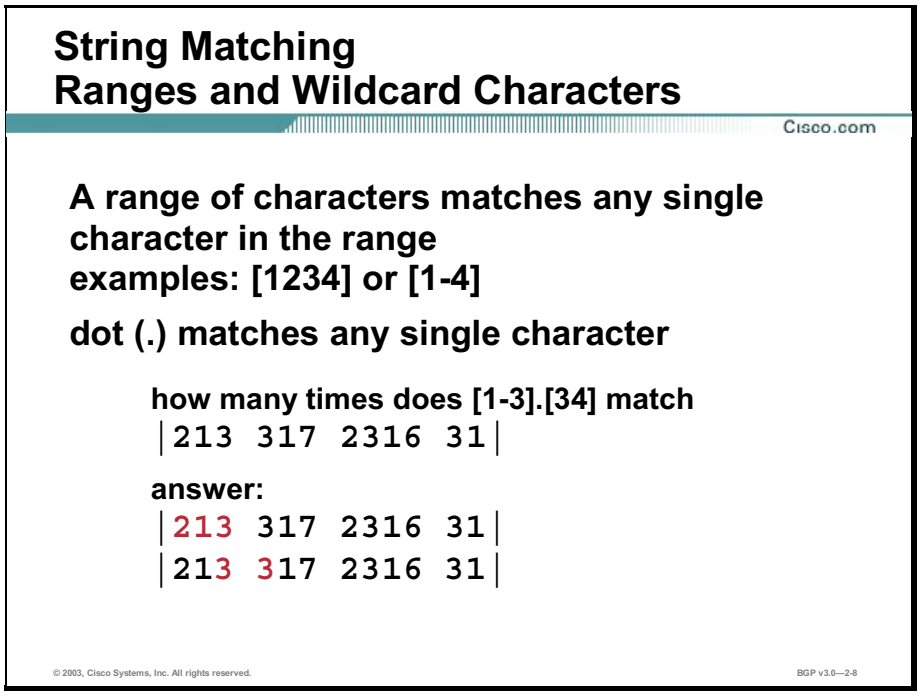
The regular expression “31” matches the AS-path string “213 *31*7 2*31*6 *31*” three times, because “31” matches a part of “317”, “2316”, and “31”.



The character "|," vertical bar, has a special meaning. It is an operator that means "or." The regular expression "21|31" matches the sequence of "2" followed by "1" or the sequence of "3" followed by "1". So, this sample regular expression will match a two-character sequence: the "21" or the "31".

The regular expression "21|31" matches the AS-path string "*21*3 *31*7 2*31*6 *31*" four times, because "21" matches a part of "213" and "31" matches a part of both "317" and "2316" as well as "31".

 Copyright © 2003, Cisco Systems, Inc.

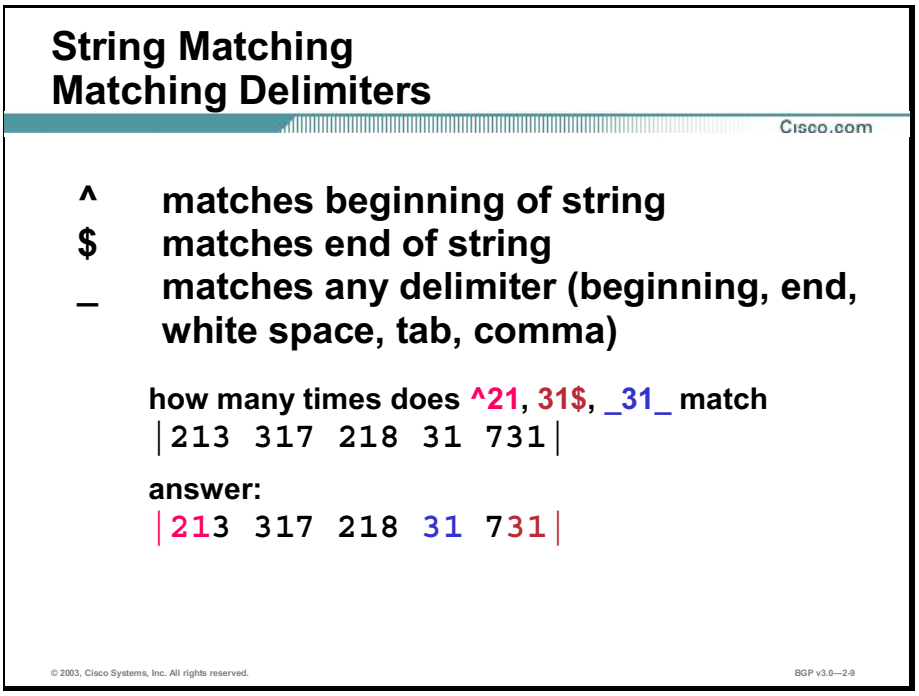

String Matching
Ranges and Wildcard Characters
Cisco.com
A range of characters matches any single character in the range
examples: [1234] or [1-4]
dot (.) matches any single character
how many times does [1-3].[34] match
|213 317 2316 31|
answer:
|213 317 2316 31|
|213 317 2316 31|
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-8

The pair of brackets "[" and "]" has a special meaning. They surround a set of characters of which any one matches. The set of characters is either expressed as the list of them (for example, "[1234]") or the sequence with the starting character, a hyphen, and the ending character (for example, "[1-4]"). Both examples match one single character, which must be any one in the set of the four characters "1", "2", "3", and "4".

The character ".", dot, matches any single character. Small regular expressions can be combined into a larger expression. Such a combination is matching if all of the parts match one after the other. The sample regular expression "[1-3].[34]" matches a sequence of three characters, of which the first must be either "1", "2", or "3", the second character can be any character, and the third must be either "3" or "4".

Note The space character delimiting two AS numbers is just a character. The dot, ".", for example, matches it.

The regular expression "[1-3].[34]" matches the AS-path string "213 317 2316 31" twice. Initially, it matches "213". The leading "[1-3]" matches the leading "2". The dot, which matches any character, matches the "1," and "[34]" matches the trailing "3". Secondly, the regular expression also matches in "21*3 3*17 2316 31". This is a little harder to see, because the dot, ".", matches the space character between "213" and "317".


String Matching
Matching Delimiters
Cisco.com
^ matches beginning of string
$ matches end of string
_ matches any delimiter (beginning, end, white space, tab, comma)
how many times does ^21, 31$, _31_ match
|213 317 218 31 731|
answer:
|213 317 218 31 731|
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-9

A character string must have a start and an end. The character with the special meaning "^" matches the beginning of a string. Because all strings have a beginning, it matches all strings. However, it is used to position the following part of the regular expression. The character following the "^" character must be the first character of the string; otherwise, it would not match the beginning of the string.

The special character "$" is used analogously, but it means the end of the string. The character preceding the "$" must be the last character in the string; otherwise, the "$" does not match the end of string.

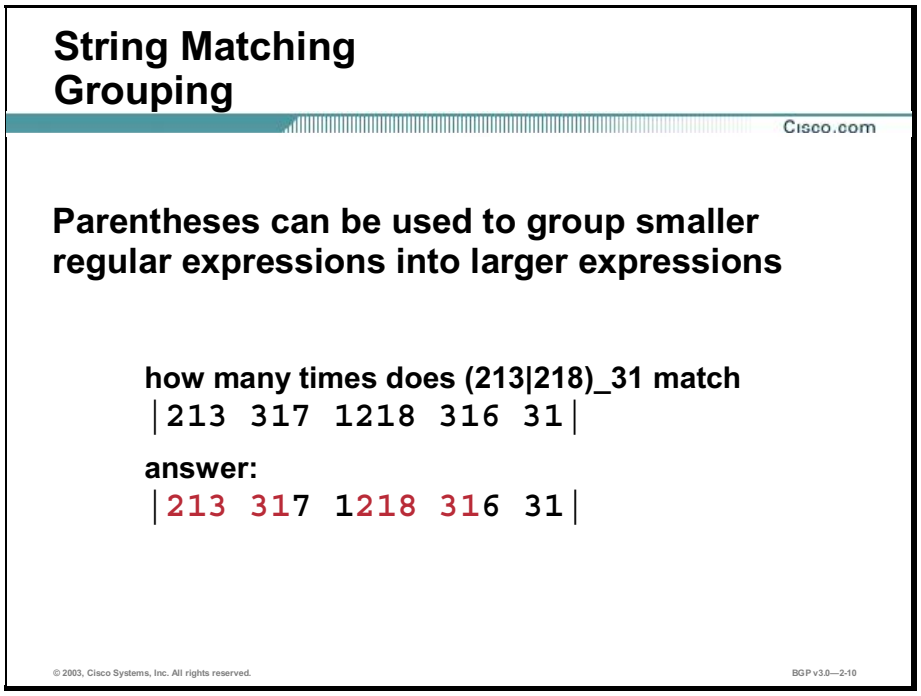
Underscore, "_", matches any delimiter. The space character between two AS numbers is an example of a delimiter. The beginning of the string and the end of the string are also considered delimiters. Underscore, "_", is used to make sure that the desired AS number is found in an AS-path string and not an AS number where the desired number could be considered a part. For example, the regular expression "31" will match the AS number string "317," but the regular expression "_31_" will not. Both "31" and "_31_" will match the desired AS number string "31".

The regular expression "^21" can match the AS-path string "213 317 218 31 731" only one time because there is only one beginning of the string. It matches only if the string starts with the sequence "21", which it does.

The regular expression "31$" can match the AS-path string "213 317 218 31 731" only one time because there is only one end of the string. It only matches if the string ends with the sequence "31", which it does.

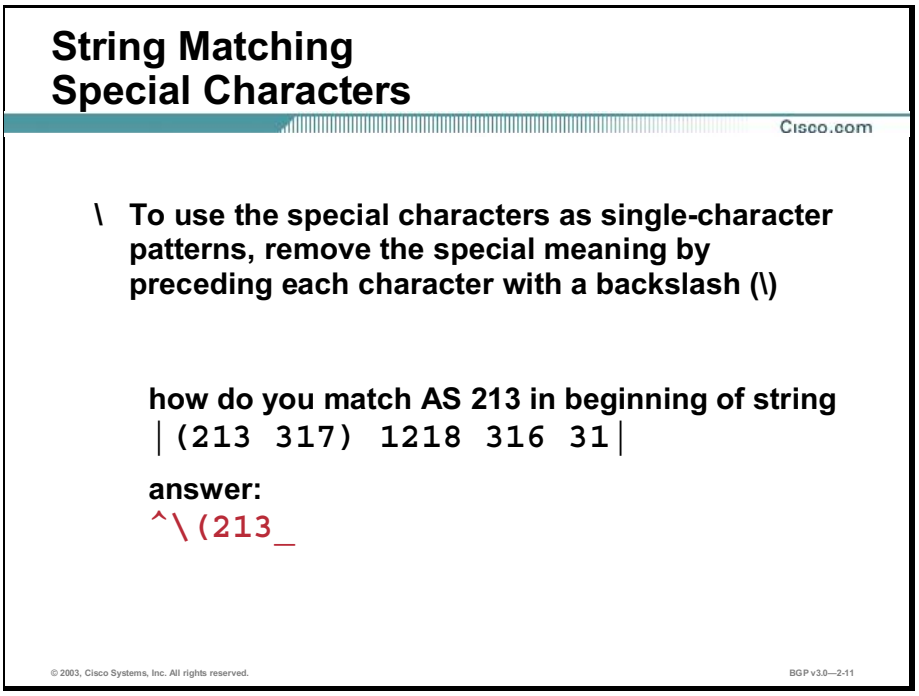
The regular expression "_31_" can, in theory, match an AS-path string several times. However, in this case, when matched against the string "213 317 218 31 731", it matches only the AS number "31" in the AS path.

 Copyright © 2003, Cisco Systems, Inc.


String Matching
Grouping
Cisco.com
Parentheses can be used to group smaller regular expressions into larger expressions
how many times does (213|218)_31 match
|213 317 1218 316 31|
answer:
|213 317 1218 316 31|
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-10

Complicated expressions must sometimes be grouped with parentheses, “(” and “)”. This feature can be useful when you are searching for a sequence of two AS numbers of which the first can be any of the two specific autonomous systems ”213” or “218”, but the last must be a third specific AS, “31”. If the brackets were not used, the expression would match either the single AS “213” or the sequence of the two “218 31”.

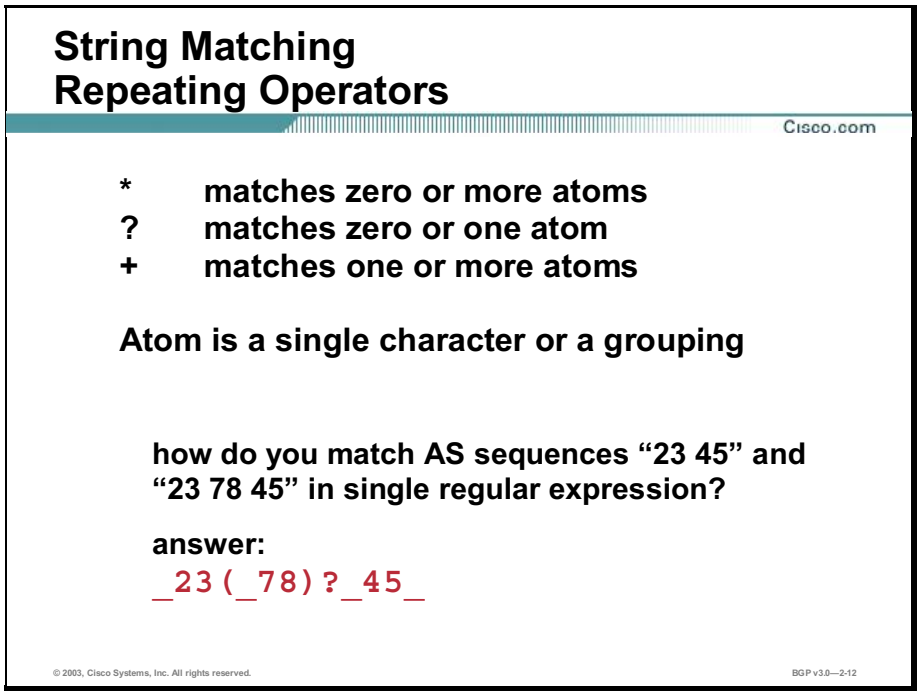
The regular expression “(213|218)_31” matches the AS-path string “213 317 1218 316 31” twice. The first match is “*213 31*7 1218 316 31”, and the second match is “213 317 1*218 31*6 31”.


String Matching
Special Characters
Cisco.com
\ To use the special characters as single-character patterns, remove the special meaning by preceding each character with a backslash (\)
how do you match AS 213 in beginning of string
|(213 317) 1218 316 31|
answer:
^\(213_
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-11

Sometimes the target string that you are trying to match with a regular expression contains some of the characters that also have special meanings in the regular expression. To match these characters in the target string, use the "\" together with the character in the regular expression.

Note	This type of regular expression syntax is used for matching AS-path strings inside a BGP confederation.

 Copyright © 2003, Cisco Systems, Inc.


String Matching
Repeating Operators
Cisco.com
* matches zero or more atoms
? matches zero or one atom
+ matches one or more atoms
Atom is a single character or a grouping
how do you match AS sequences “23 45” and “23 78 45” in single regular expression?
answer:
_23(_78)?_45_
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-12

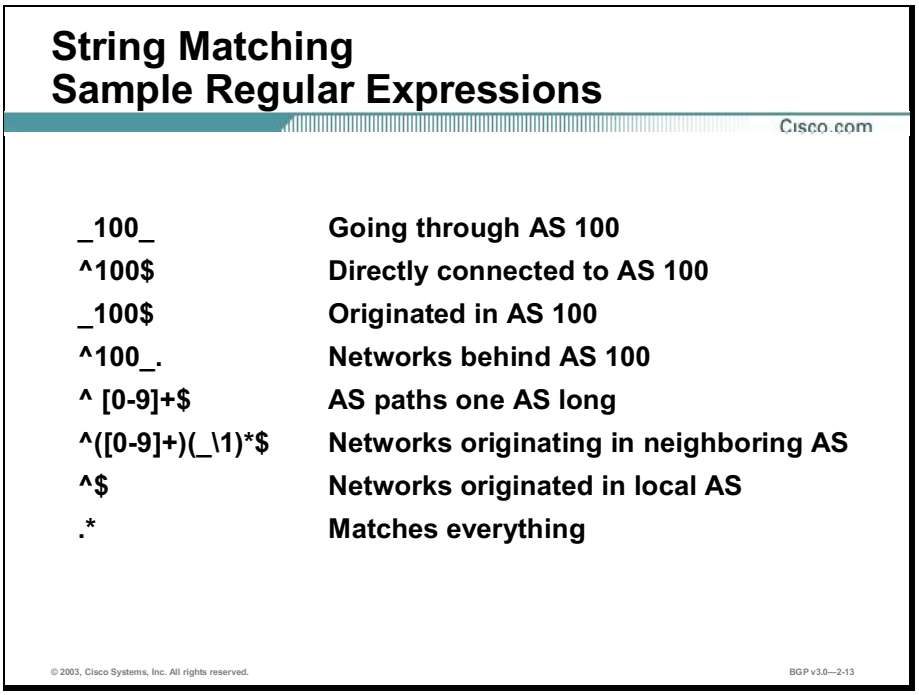
The special characters, star (asterisk), “*”, question mark, “?”, and plus, “+”, all apply repetition of the expression immediately preceding it.

Star (asterisk), “*”, means that the expression immediately preceding it is repeated zero or more times. This means that the expression may not be there, but it may also be there any number of times. The expression “1*” will match a sequence of no characters or a sequence of any number of the character “1”.

Question mark, “?”, means that the expression immediately preceding it is repeated zero or one time. This means that the expression may not be there, but it may also be there once. The expression “1?” will match a sequence of no characters or the single character “1”.

Plus, “+”, means that the expression immediately preceding it is repeated one or more times. This means that the expression must be there at least once. The expression “1+” will match a sequence of one or more of the character “1”.

Example


String Matching
Sample Regular Expressions
Cisco.com
_100_ Going through AS 100
^100$ Directly connected to AS 100
_100$ Originated in AS 100
^100_. Networks behind AS 100
^ [0-9]+$ AS paths one AS long
^([0-9]+)(_\1)*$ Networks originating in neighboring AS
^$ Networks originated in local AS
.* Matches everything
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-13

Regular expressions can be arbitrarily complex. However, most searching is accomplished using regular expressions derived from the following examples:

- If you are searching for all routes that have AS 100 in their AS paths, the regular expression to use is "_100_".
- If you are searching for all routes that are sourced in your directly connected neighboring AS 100, the regular expression to use is "^100$". Use an expression with both "^" and "$" present when you are searching for an exact match.
- If you are searching for all routes that are sourced in AS 100, but that AS is not necessarily a directly connected neighboring AS, the regular expression to use is "_100$". The "$" indicates that the AS path must end with AS 100. This is an indication that the route was sourced in AS 100. The underscore, "_", is used to make sure that it is AS 100 at the end of the string and not, for example, AS 2100.
- If you are searching for all the routes reachable behind AS 100, the regular expression to use is "^100_.". The "^" indicates that the AS path must start with AS 100. The underscore, "_", is used to make sure that it is not matching with, for example, AS 1001. The dot, ".", is used to indicate that the AS path does not end with AS 100, and that there must be something following AS 100.

 Copyright © 2003, Cisco Systems, Inc.

- If you are searching for all routes sourced in any AS directly neighboring your AS, the regular expression to use is "^[0-9]+$". The "[0-9]" part means any digit. The repetition sign plus, "+", means one or more times. Therefore, the combination "[0-9]+" means a sequence of one or more of digits. The "^" and "$" mean the beginning and the end of the string. So, the string may consist only of a sequence of one or more digits. That is one single AS number.

- If you are searching for all routes sourced in any AS directly neighboring your AS, and possibly performing AS-path prepending (multiplication of a directly connected AS number), the regular expression to use is "^([0-9]+)(_\1)*$". The expression in the first set of parentheses matches any AS number. The parentheses store the value of the matched AS, and this value is then recalled by the second part of the regular expression, including a variable. The variable "\1" is put into parentheses for the purpose of the multiplier operator "*", meaning that this part can match any number of successive occurrences of the same AS number that was matched by the "[0-9]+" expression. For example, this regular expression matches AS paths "99 99 99", "200", "101 101", or "5 5 5 5 5", but it does not match the AS path "101 99".

- The combination "^$" means an empty string and is used when searching for all routes sourced in the local AS.

- Sometimes a search is made to select a few specific routes and do something special with them, while the rest of the routes will be handled in a different way. To search for all routes, regardless of the content of their AS-path attribute, use the regular expression ".*". The dot, ".", matches any single character. The repetition character, star (asterisk), "*", means that the match should be repeated zero or more times. Thus, the combination, ".*", matches any string.

Practice

Q1) Which AS path is matched by the regular expression "72$"?

A) 213 72 218 31 727

B) 27 317 271 50 72

C) 315 27 723 19 91

D) 72 591 368 20 87

Q2) Which AS path is matched by the regular expression “$27_”?

A) 213 72 218 31 727

B) 27 317 271 50 72

C) 315 27 723 19 91

D) 72 591 368 20 27

Q3) What is the difference between the regular expressions “_100_” and “_100$”?

A) The first expression refers to routes that have the substring “100” in their AS paths; the second expression refers only to routes that are directly connected to AS 100.

B) The first expression refers to routes that have the substring “100” in their AS paths; the second expression refers only to routes that originated in AS 100.

C) The first expression refers to routes that go through AS 100; the second expression refers to routes that originated in AS 100.

D) The first expression refers to routes that are directly connected to AS 100; the second expression refers to routes that originated in AS 100.

Q4) What does the regular expression “^100$” match?

A) routes that originated in AS 100

B) routes that go through AS 100

C) routes that contain “100” in the AS path

D) routes that originate in the neighboring AS 100

Q5) How do you match AS paths that contain exactly two single-digit AS numbers?

A) Use the expression “**”.

B) Use the expression “..”.

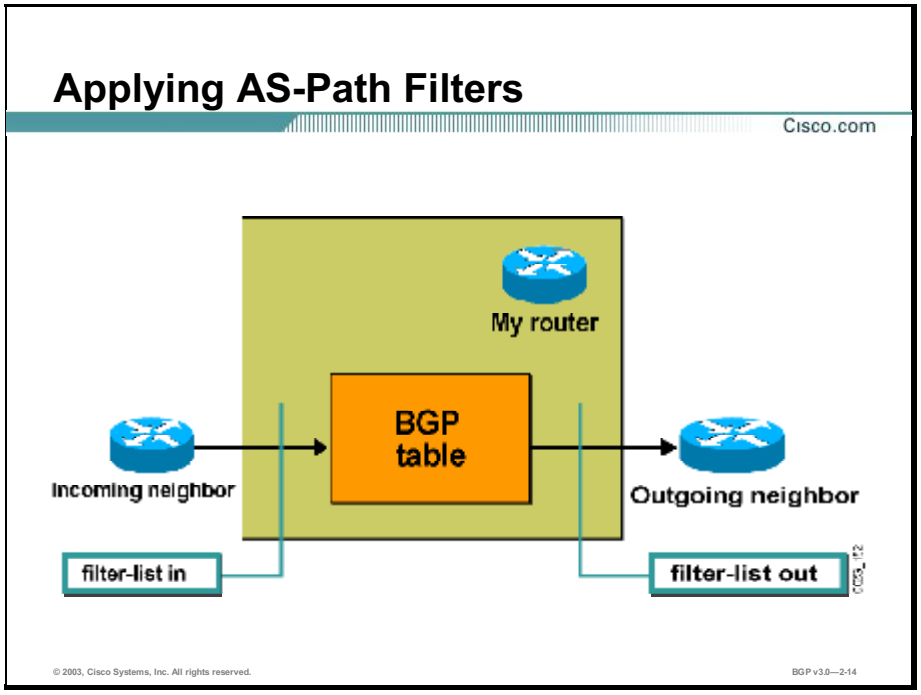
C) Use the expression “[0-9]_[0-9]”.

D) Use the expression “^[0-9]_[0-9]$”.

 Copyright © 2003, Cisco Systems, Inc.

Applying AS-Path Filters

This topic identifies where you can apply an AS-path filter when configuring a router to influence route selection.



AS-path filters configured on a router select those routes that are allowed. Routes that are selected:

- Enter the local BGP table when the selection is applied on the incoming routes from a neighbor. Routes that are not selected are silently dropped.
- Are transmitted to the neighbor when the selection is applied on the outgoing routes to the neighbor. Routes that are not selected are used locally but are never sent to the neighbor.

Practice

Q1) What happens to outgoing routes that are not selected by matching a configured AS-path filter?

A) Nonselected routes are removed from the IP routing table.

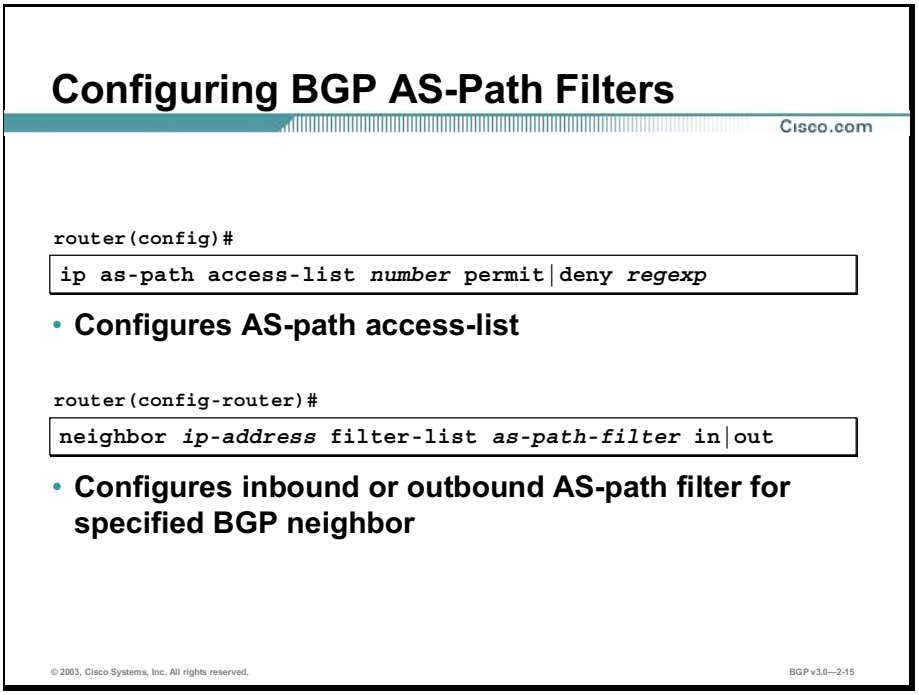
B) Nonselected routes are removed from the BGP table.

C) Nonselected routes are sent to the neighbor with a poisoned metric.

D) Nonselected routes are used by the local router only.

Configuring BGP AS-Path Filters

This topic identifies the commands to configure AS-path filters to influence route selection.


Configuring BGP AS-Path Filters
Cisco.com
router(config)#
ip as-path access-list number permit|deny regexp
• Configures AS-path access-list
router(config-router)#
neighbor ip-address filter-list as-path-filter in|out
• Configures inbound or outbound AS-path filter for specified BGP neighbor
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-15

An AS-path filter is created by an AS-path access-list. This access-list is applied to a set of routes from which to select a subset. Routes permitted by the access-list are included in the subset, and those denied are not. As in all access-lists, the candidate to be permitted or denied membership in the subset is tested against all the lines in the access-list, in the order the list is configured. The first match indicates "permit" or "deny," as specified. If the end of the access-list is reached without any explicit match, the candidate is implicitly denied.

The test by the AS-path access-list is evaluated using regular expressions applied on the AS-path attribute of the route.

The access-list can, for example, be applied on the routes received from, or those sent to, a specific BGP neighbor.

ip as-path access-list

To define a BGP AS-path access-list, use the **ip as-path access-list global** configuration command.

ip as-path access-list *access-list-number* **{permit | deny}** *as-regular-expression*

To disable use of the access-list, use the **no** form of this command.

no ip as-path access-list *access-list-number*

 Copyright © 2003, Cisco Systems, Inc.

Syntax Description

access-list-number	Integer from 1 to 199 that indicates the regular expression access-list number
permit	Permits access for matching conditions
deny	Denies access to matching conditions
as-regular-expression	AS in the access list using a regular expression. See the "Regular Expressions" appendix in the *Cisco IOS Dial Services Command Reference* for information about forming regular expressions.

neighbor filter-list

To set up a BGP filter, use the **neighbor filter-list** router configuration command.

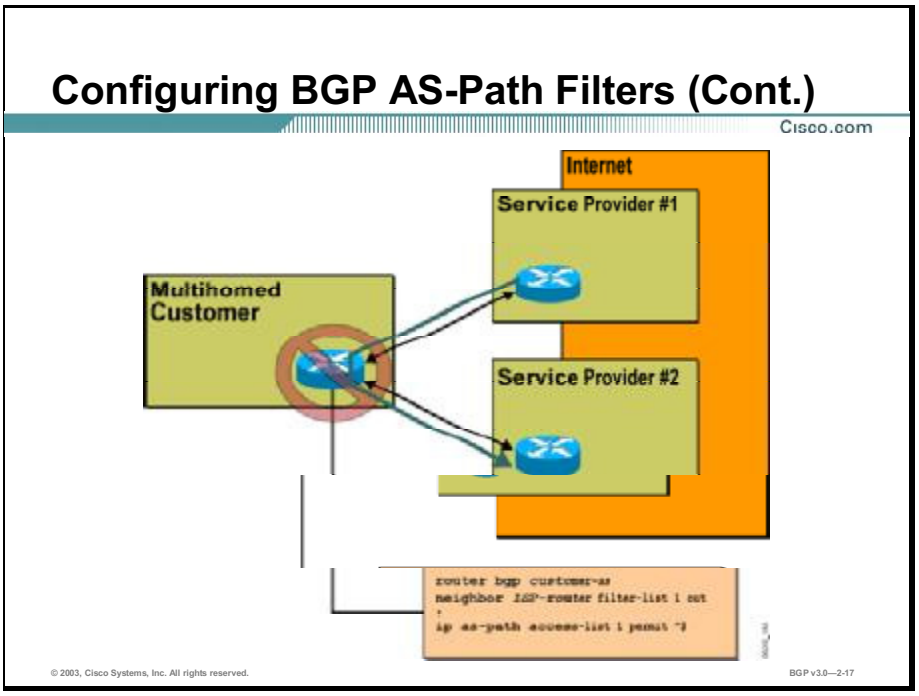
neighbor *{ip-address | peer-group-name}* **filter-list** *access-list-number* **{in | out}**

To disable this function, use the **no** form of this command.

no neighbor *{ip-address | peer-group-name}* **filter-list** *access-list-number* **{in | out}**

Syntax Description

ip-address	IP address of the neighbor.
peer-group-name	Name of a BGP peer group.
access-list-number	Number of an AS path access-list. Define this access-list with the **ip as-path access-list** command.
in	Access list to incoming routes.
out	Access list to outgoing routes.



Multihomed customers do not want to act as a transit AS between their service providers. The customer avoids this situation by not transmitting all their routes to its service providers. The service providers send IP packets to the customer only if the IP packets have destination addresses matching one of the routes that the customer has sent by BGP to the service provider. By making sure that only locally sourced routes are sent, the customer avoids receiving IP packets for destinations outside its own AS.

Within the customer AS, the locally sourced routes have empty AS paths. The empty string is matched by the regular expression "^$". The command **ip as-path access-list** *1* permits only those routes that are locally sourced and implicitly denies the rest. By applying this filter list on outgoing information to all neighbors, the customer will announce local routes only.

Practice

Q1) What three steps are required to apply a new inbound routing policy to a neighbor? (Choose three.)

A) Define an AS-path access-list.

B) Attach the AS-path filter to inbound or outbound updates for a specific BGP neighbor.

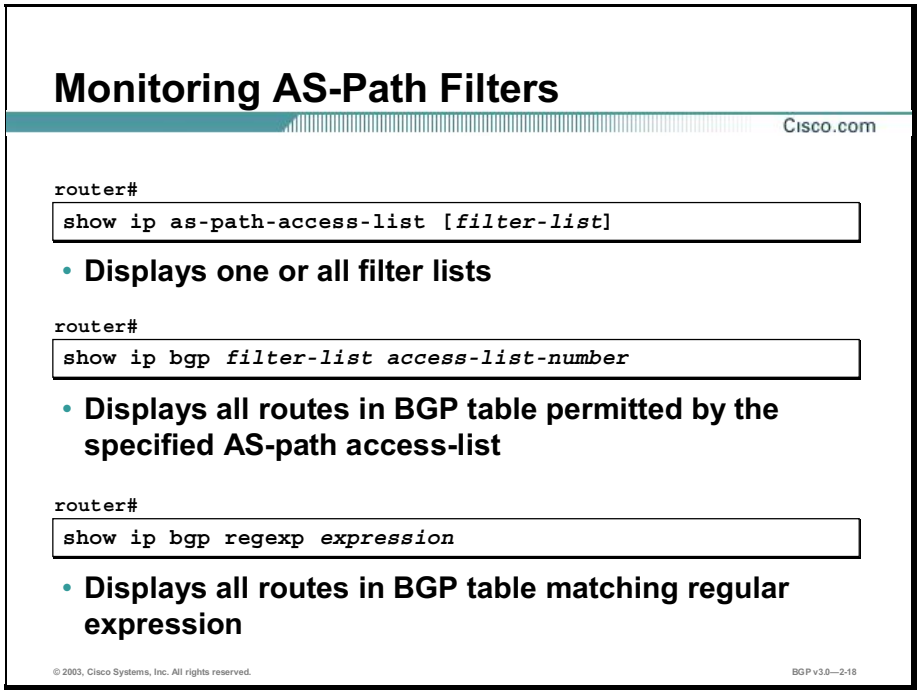
C) Send incoming and/or outgoing AS-path filters to the BGP neighbor.

D) Force the updates to go through the new filter.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring AS-Path Filters

This topic identifies the Cisco IOS commands required to monitor the operation of configured AS-path filters.


Monitoring AS-Path Filters
Cisco.com
router#
show ip as-path-access-list [filter-list]
• Displays one or all filter lists
router#
show ip bgp filter-list access-list-number
• Displays all routes in BGP table permitted by the specified AS-path access-list
router#
show ip bgp regexp expression
• Displays all routes in BGP table matching regular expression
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-18

Because regular expressions sometimes get complex, thorough testing of them is required. Use the **show ip bgp regexp** command to test a regular expression typed in on the command line. The result is a printout on the screen of all those routes currently in the BGP table that had an AS-path attribute matching the typed-in regular expression.

An AS-path access-list is even more complex because it is a combination of several regular expressions. There is one expression on each access-list line. Use the **show ip bgp filter** command to test the entire AS-path access-list. The result is a printout on the screen of all those routes currently in the BGP table that had an AS-path attribute permitted by the access-list.

show ip bgp regexp

To display routes matching the AS-path regular expression, use the **show ip bgp regexp** privileged EXEC command.

show ip bgp regexp *regular-expression*

Syntax Description

regular-expression	Regular expression to match the BGP AS paths

show ip bgp filter-list

To display routes that conform to a specified filter-list, use the **show ip bgp filter-list** privileged EXEC command.

show ip bgp filter-list *access-list-number*

Syntax Description

access-list-number	Number of an AS-path access-list. It can be a number from 1 to 199.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring AS-Path Filters (Cont.)

Cisco.com

- **Displaying configured filters**

```
router# show ip as-path-access-list
AS path access list 6
    permit ^$
AS path access list 7
    deny _213_
    permit .*
AS path access list 8
    permit _214_
AS path access list 25
    permit _42_
AS path access list 27
    deny _22_|_51$
    permit .*
```

© 2003, Cisco Systems, Inc. All rights reserved.

The **show ip as-path-access-list** displays a specific access-list or all AS-path access-lists in the router.

Monitoring AS-Path Filters (Cont.)

Cisco.com

- **Routes matched by a filter list**

```
router# show ip as-path-access-list 25
AS path access list 25
    permit _42_

router# show ip bgp filter-list 25
BGP table version is 61, local router ID is 197.6.2.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 128.37.0.0       192.168.21.2                100      0 (65002 65003 65004) 99 2 20 42 37 i
*> 128.42.0.0       192.168.21.7                100      0 (65002 65003 65004) 99 7 20 42 i
*> 192.26.11.0      192.168.20.20            0             0 20 42 26 i
*> 192.37.11.0      192.168.20.20            0             0 20 42 37 i
*> 192.42.11.0      192.168.20.20            0             0 20 42 i
```

© 2003, Cisco Systems, Inc. All rights reserved.

The AS-path access-list number 25 in this example consists of one single line. It permits those routes that have an AS-path attribute containing the AS number 42. All other routes are implicitly denied.

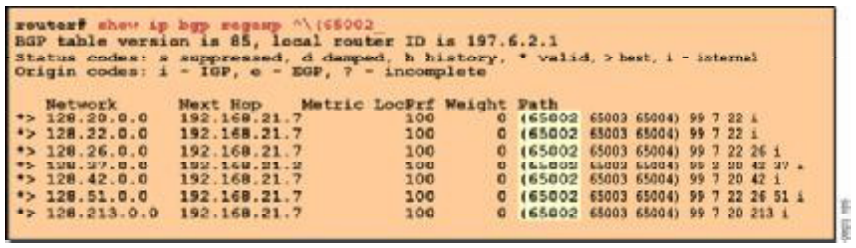
The **show ip bgp filter-list** command is used to display all the routes currently in the BGP table that are permitted by AS-path access-list 25. It will show all the routes with AS 42 somewhere in their AS paths.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring AS-Path Filters (Cont.)

Cisco.com

- **Routes matched by an expression**



© 2003, Cisco Systems, Inc. All rights reserved.

The **show ip bgp regexp** command displays all routes currently in the BGP table that have an AS-path attribute matching the typed-in regular expression. In this example, BGP confederations, a scalability feature, are in use. As a result of configuring BGP confederations, the AS path contains some AS numbers enclosed in parentheses. BGP confederations and their usage is explained later. It is the AS path regular expression that is important in this example.

In this example, you wish to find all BGP confederation routes from AS number 65002. To search for this character in the beginning of the string, use the character “\”, backslash. The regular expression “^\(65002” matches those routes received from the intraconfederation AS number 65002.

Practice

Q1) How can you test your regular expression?

A) **show ip bgp access-list** command

B) **show ip bgp filter** command

C) **show ip bgp regexp** command

D) **show ip bgp summary** command

Q2) How can you test your AS-path filter before using it?

A) **show ip bgp access-list** command

B) **show ip bgp filter** command

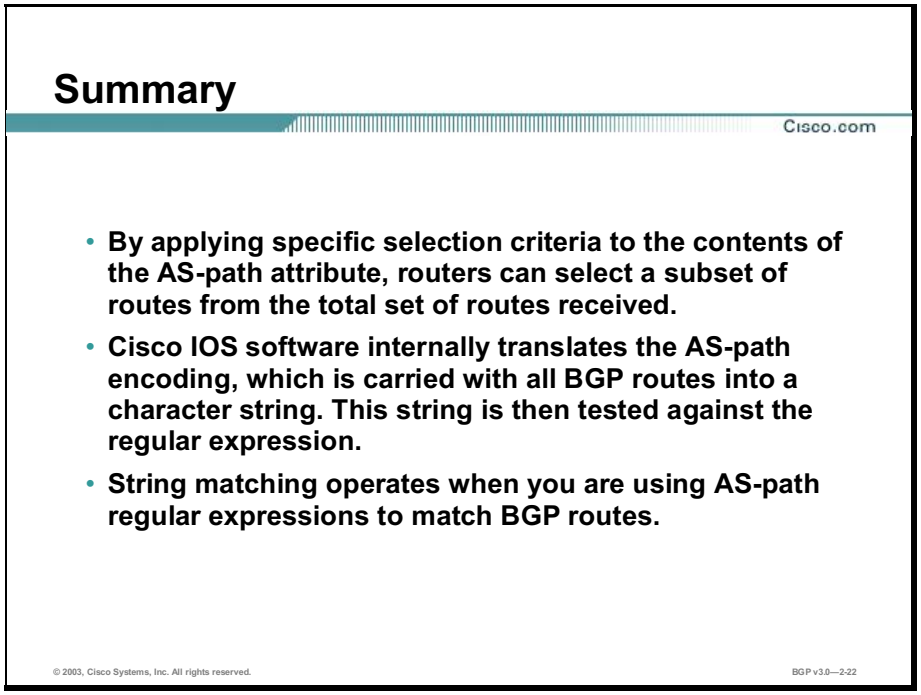
C) **show ip bgp regexp** command

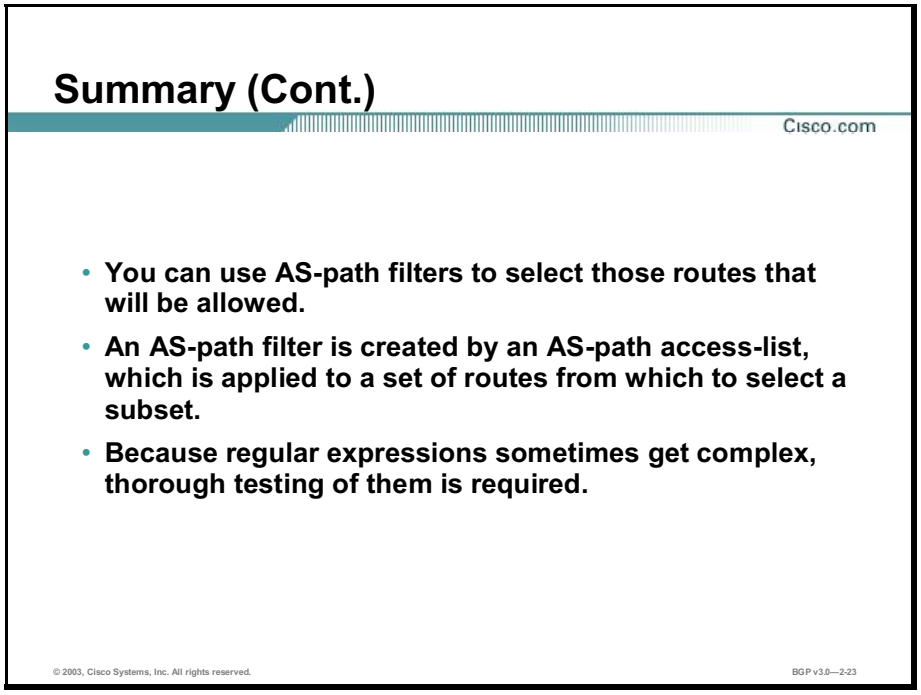
D) **show ip bgp summary** command

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• By applying specific selection criteria to the contents of the AS-path attribute, routers can select a subset of routes from the total set of routes received.
• Cisco IOS software internally translates the AS-path encoding, which is carried with all BGP routes into a character string. This string is then tested against the regular expression.
• String matching operates when you are using AS-path regular expressions to match BGP routes.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-22


Summary (Cont.)
Cisco.com
• You can use AS-path filters to select those routes that will be allowed.
• An AS-path filter is created by an AS-path access-list, which is applied to a set of routes from which to select a subset.
• Because regular expressions sometimes get complex, thorough testing of them is required.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-23

Next Steps

After completing this lesson, go to:

- Prefix-List Filters lesson

References

For additional information, refer to these resources:

- For more information on regular expressions, refer to “Using Regular Expressions in BGP” at the following URL: http://www.cisco.com/warp/public/459/26.html
- For more information on AS-path filters, refer to “BGP Case Studies Section 3” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: AS-Path Filters

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using AS-path filters where connections to multiple ISPs must be supported.

After completing this exercise, you will be able to:

- Identify where you can apply an AS-path filter when configuring a router to influence route selection
- Configure AS-path filters to influence route selection
- Monitor the operation of configured AS-path filters

Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor *ip-address* **filter-list** *n* {**in**|**out**}	Filter incoming or outgoing updates with the specified AS-path filter.
ip as-path access-list *n* **permit regexp**	Specify the AS-path filter.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.
show ip bgp filter-list *n*	Use this command to check a filter.
clear ip bgp *neighbor*	Clears the BGP session.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- Your service providers complain that you propagate routes between them and that you could potentially become a transit network.
- You have also discovered that the service provider "Cheap" does not provide you with good connectivity to your users in AS 213. Therefore, you want to accept the routes originating in AS 213 only from the provider "Good."
- Customers in AS 214 request that you use only the provider "Cheap" to communicate with them due to security reasons.
- Use AS-path filters to filter incoming routes and to limit propagation of transit routes.
- Select routes between the "Good" and "Cheap" service providers based on the following criteria:
  - Accept routes originating in AS 213 only from router "Good"
  - Accept routes with AS 214 in the AS path only from router "Cheap"
  - Ensure that you do not act as a transit AS between the two service provider routers ("Good" and "Cheap").
- Figure 1 shows the structure of the autonomous systems beyond AS 20 and AS 22. This information will assist you in your verification and any troubleshooting steps.

 Copyright © 2003, Cisco Systems, Inc.

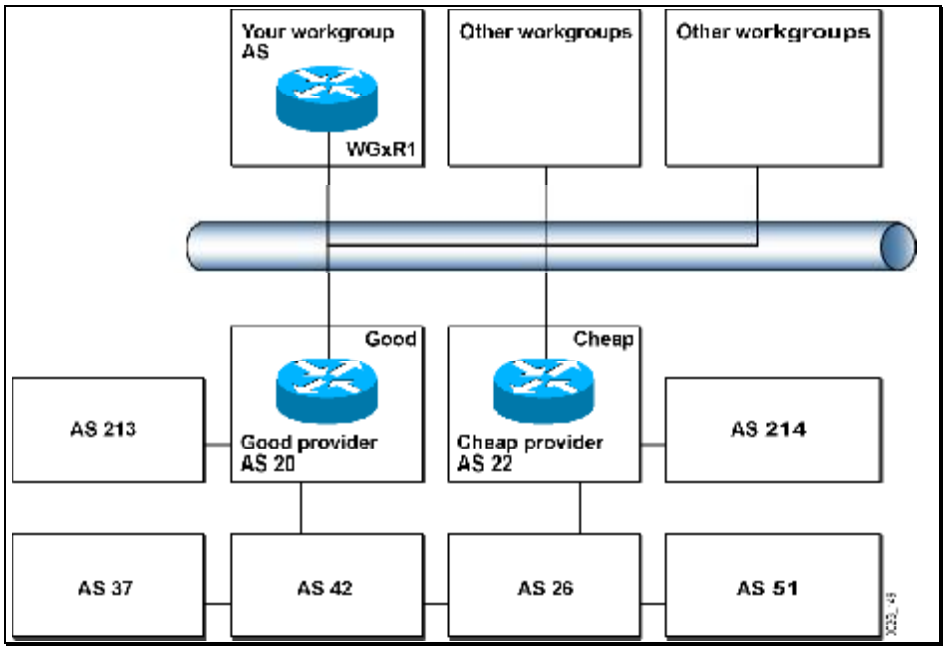


Figure 1: Structure of autonomous systems beyond your upstream service providers

Exercise Procedure

Complete these steps:

Step 1 Log on to the “Good” and “Cheap” routers and verify the service provider claims that you propagate the routes between them.

Router “Good”:

```
   Network          Next Hop            Metric LocPrf Weight Path
*  192.168.1.0      192.168.20.1                           0 22 1 i
*>                  192.168.20.1                0          0 1 i
*  197.1.0.0/21     192.168.20.1                           0 22 1 i
*>                  192.168.20.1                           0 1 i
*  197.1.0.0/16     192.168.20.1                           0 22 1 i
*>                  192.168.20.1                           0 1 i
*  197.1.8.0/22     192.168.20.1                           0 22 1 i
*>                  192.168.20.1                           0 1 i
```

Router “Cheap”:

```
   Network          Next Hop            Metric LocPrf Weight Path
*> 192.168.1.0      192.168.20.1                0          0 1 i
*                   192.168.20.1                           0 20 1 i
*> 197.1.0.0/21     192.168.20.1                           0 1 i
*                   192.168.20.1                           0 20 1 i
*> 197.1.0.0/16     192.168.20.1                           0 1 i
*                   192.168.20.1                           0 20 1 i
*> 197.1.8.0/22     192.168.20.1                           0 1 i
*                   192.168.20.1                           0 20 1 i
```

Configuring AS-path access–lists:

Step 2 With the **ip as-path access-list** command, configure an AS-path access-list to deny routes that have number “213” as the last number in the AS path. The filter should permit all other routes.

Step 3 Configure another AS-path access-list to deny routes that contain number 214 in the AS path. The filter should permit all other routes.

Step 4 Configure an AS-path access-list to permit prefixes originating in your AS and deny all other prefixes.

BGP table showing prefixes that have “213” or “214” in their path (before applying the AS-path filters):

Before AS-path filters:

```
wg1r1#sh ip bgp
BGP table version is 122, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric LocPrf Weight Path
*> 128.213.0.0      192.168.20.20                         100 22 20
213 i
*                   192.168.20.20            0             0 20 213 i
*> 128.214.0.0      192.168.20.22            0           100 22 214 i
*> 192.213.11.0     192.168.20.20                         100 22 20
213 i
*                   192.168.20.20            0             0 20 213 i
*> 192.214.11.0     192.168.20.22            0           100 22 214 i
*                   192.168.20.22            0             0 22 214 i
```

Applying AS-path access–lists:

Step 5 Apply the AS-path access-lists configured in the previous steps to your BGP neighbors.

Exercise Verification

You have completed this exercise when you attain these results:

- Verify that all paths containing AS 213 are received only from router “Good” and that all paths containing AS 214 are received only from router “Cheap.”

```
WG1R1#sh ip bgp
BGP table version is 84, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  128.20.0.0       192.168.20.20            0             20 i
*>                  192.168.20.22            0             22 i
*  128.22.0.0       192.168.20.20            0             20 i
*>                  192.168.20.22            0             22 i
*  128.26.0.0       192.168.20.20            0             20 42 26
i
*>                  192.168.20.22            0             22 26 i
*  128.37.0.0       192.168.20.20            0             20 42 37
i
```

 Copyright © 2003, Cisco Systems, Inc.

```
*>                  192.168.20.22           0              22 26 42
37 i
*  128.42.0.0       192.168.20.20           0              20 42 i
*>                  192.168.20.22           0              22 26 42
i
*  128.51.0.0       192.168.20.20           0              20 42 26
51 i
*>                  192.168.20.22           0              22 26 51
i
*> 128.213.0.0      192.168.20.20           0              20 213 i
*> 128.214.0.0      192.168.20.22           0              22 214 i
*  192.20.11.0      192.168.20.20           0              20 i
*>                  192.168.20.22           0              22 i
*  192.22.11.0      192.168.20.20           0              20 i
*>                  192.168.20.22           0              22 i
*  192.26.11.0      192.168.20.20           0              20 42 26
i
*>                  192.168.20.22           0              22 26 i
*  192.37.11.0      192.168.20.20           0              20 42 37
i
*>                  192.168.20.22           0              22 26 42
37 i
*  192.42.11.0      192.168.20.20           0              20 42 i
*>                  192.168.20.22           0              22 26 42
i
*  192.51.11.0      192.168.20.20           0              20 42 26
51 i
*>                  192.168.20.22           0              22 26 51
i
. . .
*  192.213.11.0     192.168.20.20           0              20 213 i
*> 192.214.11.0     192.168.20.22           0              22 214 I
. . .
WG1R1#
```

- Verify that routers “Good” and “Cheap” are not receiving any prefixes that do not originate in your AS.

```
Good> sh ip bgp reg ^1_
BGP table version is 47, local router ID is 199.199.199.199
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 192.168.1.0/30   192.168.20.1             0             0 1 i
*> 192.168.1.0      192.168.20.1             0             0 1 i
*> 192.168.1.4/30   192.168.20.1       2681856             0 1 i
*> 192.168.1.8/30   192.168.20.1       3193856             0 1 i
*> 197.1.0.0/21     192.168.20.1                           0 1 i
*> 197.1.0.0/16     192.168.20.1                           0 1 i
*> 197.1.8.0/22     192.168.20.1                           0 1 i
Good>
```

Answer these questions:

Q1) By this time, the BGP table should contain a large number of prefixes. What regular expression would you use with the **show ip bgp** command on WGxR1 to view prefixes originated by your AS?

Q2) What regular expression would you use on router “Good” or “Cheap” to view prefixes originating in your AS?

Q3) What regular expression would you use on router “Good” or “Cheap” to view prefixes received from your AS?

Q4) How can you test your AS-path filters before applying them and clearing BGP neighbors?

 Copyright © 2003, Cisco Systems, Inc.

Prefix-List Filters

Overview

Influencing route selection or enforcing administrative policies with packet filtering is common in network designs where customers connect to the Internet. In multihomed customer implementations where Border Gateway Protocol (BGP) is used, using access-lists to filter the large number of advertised routes can create performance bottlenecks. Prefix-lists offer a significant performance improvement over access-list filtering in loading and route lookup in large lists.

This lesson discusses the requirement for using prefix-based filters in customer implementations where connections to multiple Internet service providers (ISPs) must be supported, and discusses the benefits of using prefix-lists versus IP access-lists. The commands to apply filtering of inbound or outbound updates with prefix-lists and to configure prefix-list filters will be discussed as well as where network administrators should apply them.

Importance

Where multiple paths between a customer and ISP exist, there is a requirement to filter certain information during BGP updates to influence route selection or to enforce an administrative policy. To meet this requirement, you must use filters. There are requirements for using prefix-based filters in these circumstances. Using prefix-lists is typically easier than using standard IP access–lists, and provides performance benefits. It is important to understand the commands to apply filtering of inbound or outbound updates with prefix-lists and where they should be applied.

Objectives

Upon completing this lesson, you will be able to:

- Identify the requirement for prefix-based filters in customer implementations where you must support connections to multiple ISPs

- List the benefits of prefix-lists versus IP access-lists and describe the applications of prefix-lists in BGP networks
- Identify the Cisco IOS® commands required to configure prefix-list filters
- Describe where you can implement prefix-lists in a BGP network
- Identify the Cisco IOS commands required to apply filtering of inbound or outbound updates with prefix-lists
- Identify the Cisco IOS commands required to modify configured prefix-list filters
- Identify the Cisco IOS commands required to monitor the operations of configured prefix-list filters

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

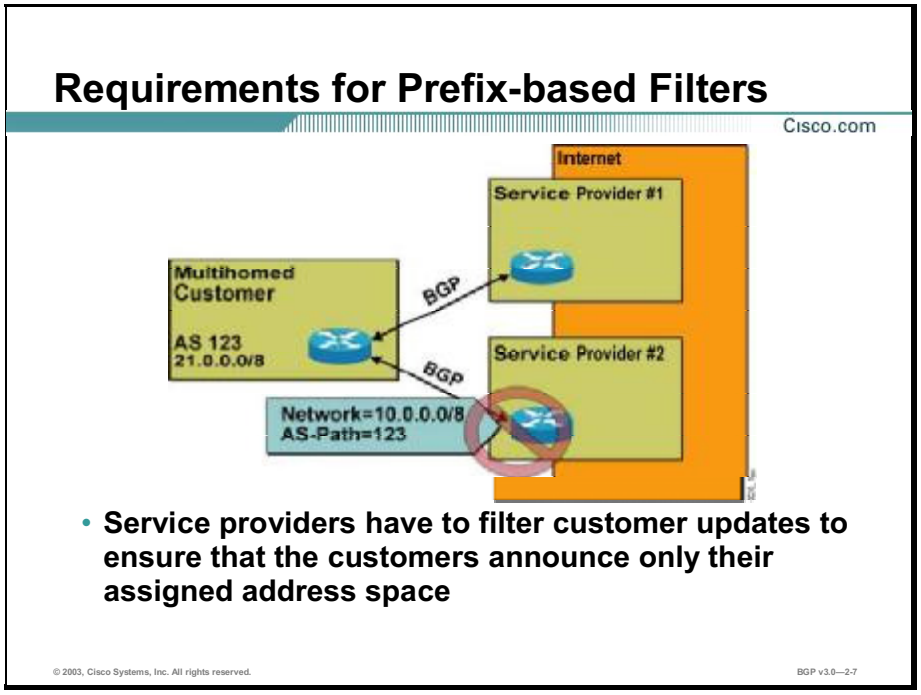
This lesson includes these topics:

- Overview
- Requirements for Prefix-Based Filters
- Prefix-Lists vs. IP Access-Lists
- Configuring Prefix-Lists
- BGP Filters Implementation
- Implementing Prefix-Lists in the BGP Process
- Modifying Prefix-Lists
- Monitoring Prefix-Lists
- Summary
- Assessment (Lab): Prefix-List Filters

 Copyright © 2003, Cisco Systems, Inc.

Requirements for Prefix-Based Filters

This topic identifies the requirement for prefix-based filters in network implementations where multiple connections between a customer and ISPs exist.



Customers with multihomed networks are responsible for announcing their own networks using BGP. Typically, customers are not as experienced with BGP as service providers, and therefore problems are more likely to occur. A service provider with a multihomed customer must take precautions not to accept, use, or forward any erroneous routing information received from the customer.

The customer is assigned a set of IP network numbers that it should announce. If the customer announces any additional networks, something is wrong. The customer may have forgotten not to act as a transit and has started propagating routes that it has received from the other service provider. Or, the customer has accidentally started to announce its private address space, which the customer may use for address links, loopback interfaces, or other devices that should never access the Internet.

To avoid problems, the service provider can apply an IP prefix filter on the incoming information from the customer. The service provider will accept only network numbers permitted by an access-list or prefix-list.

Practice

Q1) What are two reasons that a multihomed customer needs prefix-lists? (Choose two.)

A) to ensure that only valid IP prefixes are announced to the ISPs

B) to set a limit on the number of prefixes that it can accept from the ISPs

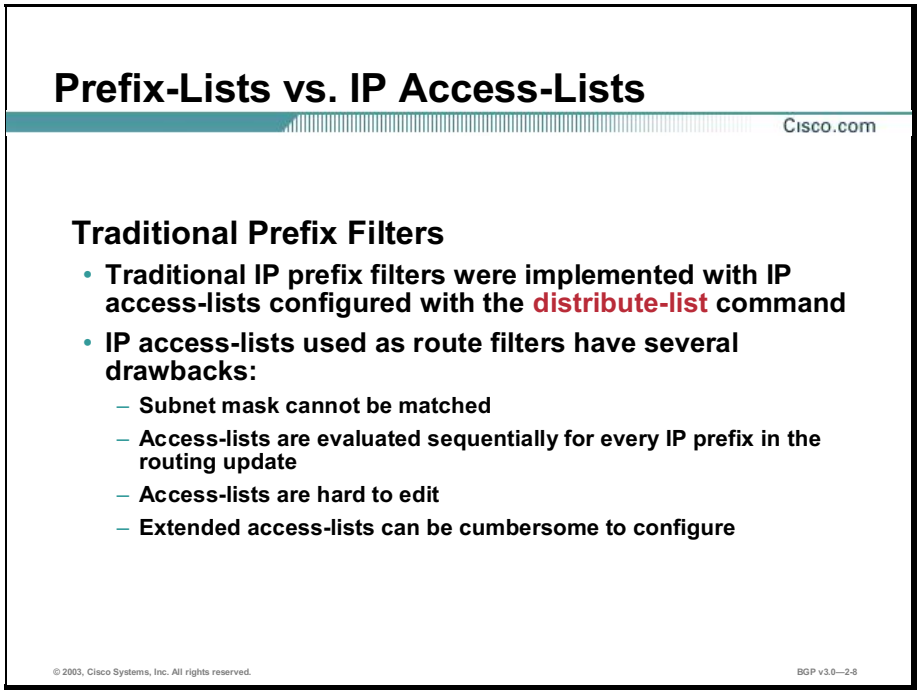
C) to prevent the customer from receiving its own IP prefixes from the ISP

D) to verify that the customer has received full Internet route tables

 Copyright © 2003, Cisco Systems, Inc.

Prefix-Lists vs. IP Access-Lists

This topic lists the benefits of prefix-lists versus IP access-lists and describes the applications of prefix-lists in BGP networks.

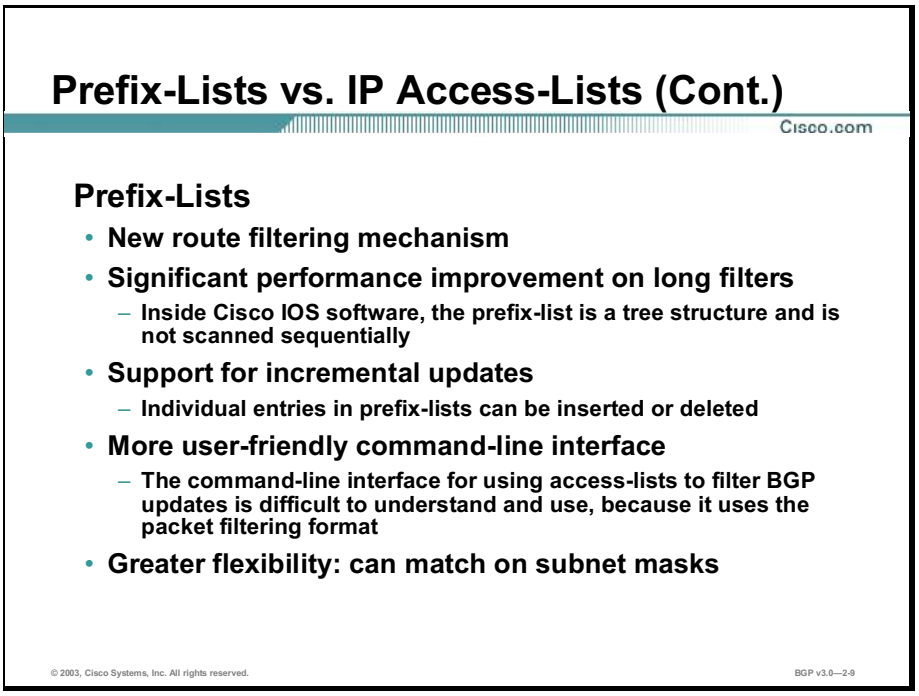

Prefix-Lists vs. IP Access-Lists
Cisco.com
Traditional Prefix Filters
• Traditional IP prefix filters were implemented with IP access-lists configured with the distribute-list command
• IP access-lists used as route filters have several drawbacks:
– Subnet mask cannot be matched
– Access-lists are evaluated sequentially for every IP prefix in the routing update
– Access-lists are hard to edit
– Extended access-lists can be cumbersome to configure
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-8

Traditionally, filtering of IP network numbers has been accomplished using an access-list. The access-list is then bound to either incoming or outgoing information of a neighbor by the **neighbor distribute-list** command. A BGP update about a network number that is permitted by the access-list will be accepted, and those denied will be dropped.

However, standard access-lists do not support testing of the subnet masks. If the access-list permits 10.0.0.0/16, it would also permit 10.0.0.0/8.

Extended access-lists can do testing on both an IP network number and a subnet mask, but the syntax is cumbersome.

Lastly, access-lists are hard to edit. The router will automatically add new access list entries to the end of the list. Because the router evaluates the list sequentially, and the first match results in a “permit” or “deny,” the order of the lines in the access-list is of utmost importance. Therefore, the inability to add a line in the middle of a list has been an administrative burden.


Prefix-Lists vs. IP Access-Lists (Cont.)
Cisco.com
Prefix-Lists
• New route filtering mechanism
• Significant performance improvement on long filters
– Inside Cisco IOS software, the prefix-list is a tree structure and is not scanned sequentially
• Support for incremental updates
– Individual entries in prefix-lists can be inserted or deleted
• More user-friendly command-line interface
– The command-line interface for using access-lists to filter BGP updates is difficult to understand and use, because it uses the packet filtering format
• Greater flexibility: can match on subnet masks
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-9

The **ip prefix-list** configuration command has several benefits compared to using the **access-list** command. The intended use of prefix-lists is limited to route filtering, whereas access-lists were initially intended for packet filtering, which was then extended to filter routes.

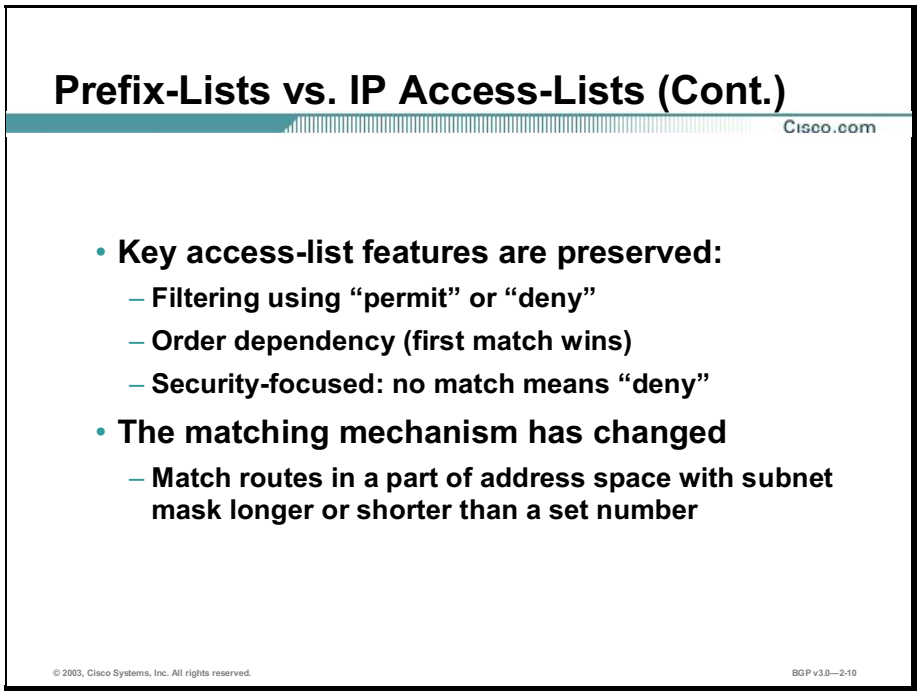
The prefix-list is internally transformed into a tree-structure, with each branching of the tree a test. The Cisco IOS software determines a verdict of either “permit” or “deny” much faster this way, compared to sequentially interpreting an access-list.

The configuration command line interface (CLI) that you use when configuring the **ip prefix-list** provides the ability to assign a line number to each line of the prefix-list. The router will use this number to sort the entries in the prefix-list. If the lines are initially assigned line numbers, with some spacing in between them, administrators can add additional lines at a later time and insert them where required. Individual lines can also be removed without removing the entire list.

Routers match network numbers in a routing update against the prefix-list using as many bits as indicated. For example, a prefix-list can be specified to be 10.0.0.0/16, which will match 10.0.0.0 routes but not 10.1.0.0 routes.

Optionally, the prefix-list can also specify the size of the subnet mask. In addition, the prefix-list can indicate that the subnet mask must be in a specified range.

 Copyright © 2003, Cisco Systems, Inc.


Prefix-Lists vs. IP Access-Lists (Cont.)
Cisco.com
• Key access-list features are preserved:
– Filtering using “permit” or “deny”
– Order dependency (first match wins)
– Security-focused: no match means “deny”
• The matching mechanism has changed
– Match routes in a part of address space with subnet mask longer or shorter than a set number
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-10

The prefix-list has several similarities with the access-list. It can consist of any number of lines, each of which indicates a test and a result. The router can interpret the lines in the specified order, although this is optimized in the Cisco IOS software. When a router evaluates a route against the prefix-list, the first line that matches will result in either a “permit” or “deny.” If none of the lines in the list match, the result is “implicitly deny.”

Testing is done using prefixes. The indicated number of bits in the prefix is compared with the same number of bits in the network number in the update. If they match, testing continues by testing the number of bits set in the subnet mask. The prefix-list line can indicate a range in which the number must be to pass the test. If no range is indicated, the subnet mask must match the prefix size.

Practice

Q1) What are three benefits of prefix-lists when compared to IP access-lists? (Choose three.)

A) Prefix-lists have a tree structure so the “permit” or “deny” result can be reached faster.

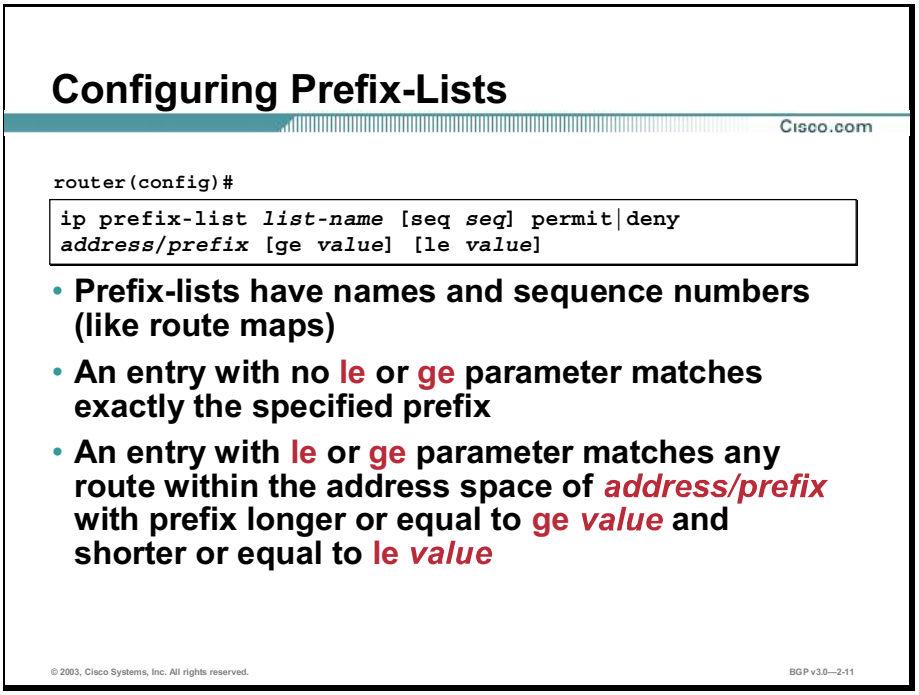
B) Prefix-lists can indicate that the subnet mask must be within a specified range.

C) If none of the lines in a prefix-list match, the result is “implicitly deny.”

D) Individual entries in prefix-lists are easier to insert or delete.

Configuring Prefix-Lists

This topic identifies the Cisco IOS commands required to configure prefix-list filters.


Configuring Prefix-Lists
Cisco.com
router(config)#
ip prefix-list list-name [seq seq] permit|deny
address/prefix [ge value] [le value]
• Prefix-lists have names and sequence numbers (like route maps)
• An entry with no le or ge parameter matches exactly the specified prefix
• An entry with le or ge parameter matches any route within the address space of address/prefix with prefix longer or equal to ge value and shorter or equal to le value
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-11

ip prefix-list

To create an entry in a prefix-list, use the **ip prefix-list** global configuration command. To delete the entry, use the **no** form of this command.

ip prefix-list *list-name* [**seq** *seq-value*] **deny** | **permit** *network/len* [**ge** *ge-value*] [**le** *le-value*]

no ip prefix-list *list-name* [**seq** *seq-value*] **deny** | **permit** *network/len* [**ge** *ge-value*] [**le** *le-value*]

Syntax Description

list-name	Name of a prefix-list
seq	(Optional) Applies the sequence number to the prefix-list entry being created or deleted
seq-value	(Optional) Specifies the sequence number for the prefix-list entry
deny	Denies access to matching conditions
permit	Permits access for matching conditions
network/len	(Mandatory) The network number and length (in bits) of the network mask
ge	(Optional) Applies the *ge-value* to the range specified

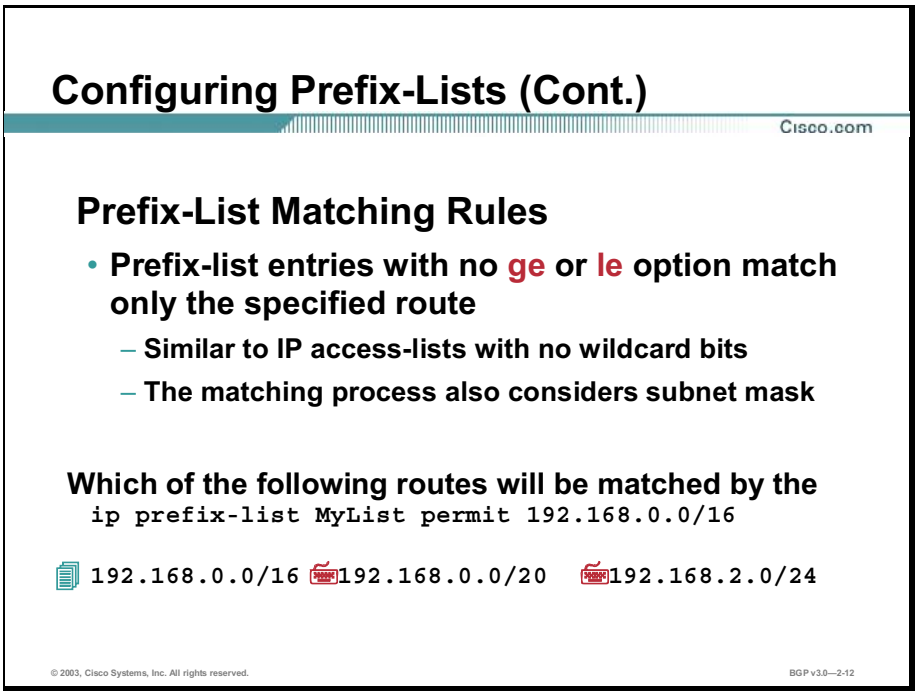
 Copyright © 2003, Cisco Systems, Inc.

ge-value	(Optional) Specifies the lesser value of a range (the "from" portion of the range description)
le	(Optional) Applies the *le-value* to the range specified
le-value	(Optional) Specifies the greater value of a range (the "to" portion of the range description)

When multiple entries of a prefix-list match a given prefix, the sequence number of a prefix-list entry identifies the entry with the lowest sequence number. In this case, the entry with the smallest sequence number is considered to be the “real” match.

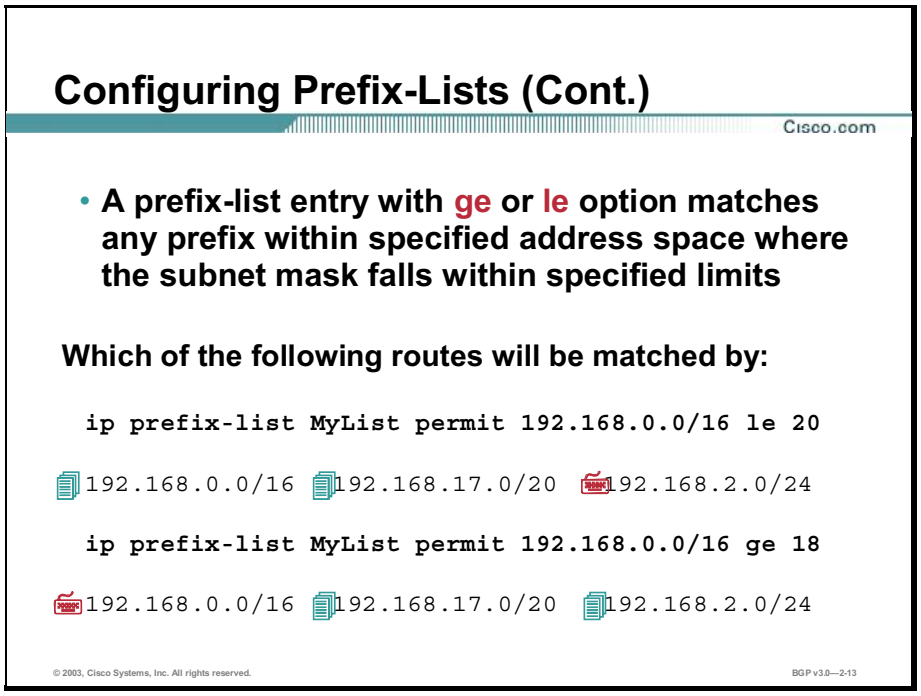
Note	You can specify sequence values for prefix-list entries in any increments you want (the automatically generated numbers are incremented in units of 5). If you specify the sequence values in increments of 1, you will not be able to insert additional entries into the prefix-list. If you choose very large increments, you could run out of sequence values.

You can use the parameters **ge** and **le** to specify the range of the prefix length to be matched for prefixes that are more specific than *network/len*. The exact match is assumed when neither **ge** nor **le** is specified. The range is assumed to be from *ge-value* to 32 only if the **ge** attribute is specified. The range is assumed to be from *len* to *le-value* only if the **le** attribute is specified.


Configuring Prefix-Lists (Cont.)
Cisco.com
Prefix-List Matching Rules
• Prefix-list entries with no ge or le option match only the specified route
– Similar to IP access-lists with no wildcard bits
– The matching process also considers subnet mask
Which of the following routes will be matched by the
ip prefix-list MyList permit 192.168.0.0/16
192.168.0.0/16
192.168.0.0/20
192.168.2.0/24
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-12

Prefix-list entries without the **ge** or **le** option match only the route with the specified IP address and subnet mask. In the example above, the prefix-list entry **permit 192.168.0.0/16** will not match the route 192.168.2.0/24 because of the mismatch in the IP address. It will also not match the route 192.168.0.0/20 because of the mismatch in the subnet mask.

 Copyright © 2003, Cisco Systems, Inc.

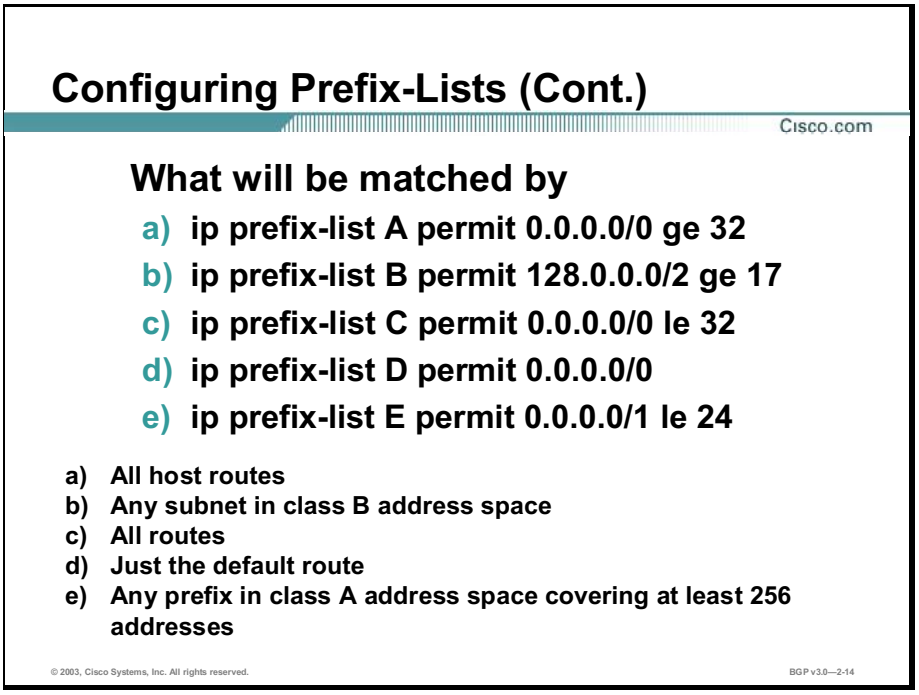

Configuring Prefix-Lists (Cont.)
Cisco.com
• A prefix-list entry with ge or le option matches any prefix within specified address space where the subnet mask falls within specified limits
Which of the following routes will be matched by:
ip prefix-list MyList permit 192.168.0.0/16 le 20
192.168.0.0/16
192.168.17.0/20
192.168.2.0/24
ip prefix-list MyList permit 192.168.0.0/16 ge 18
192.168.0.0/16
192.168.17.0/20
192.168.2.0/24
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-13

Prefix-list entries with the **ge** or **le** option specified match any prefix within the address space specified with the *address/prefix* parameter, as long as the subnet mask length of the route falls within the range specified by the **le** and **ge** parameters.

In the first example above, the route 192.168.2.0/24 is not matched by prefix-list entry **permit 192.168.0.0/16** even though the IP address falls within the specified address range, because the subnet mask is too long.

In the second example, the route 192.168.0.0/16 is not matched by prefix-list entry **permit 192.168.0.0/18** because the subnet mask is too short.

Example


Configuring Prefix-Lists (Cont.)
Cisco.com
What will be matched by
a) ip prefix-list A permit 0.0.0.0/0 ge 32
b) ip prefix-list B permit 128.0.0.0/2 ge 17
c) ip prefix-list C permit 0.0.0.0/0 le 32
d) ip prefix-list D permit 0.0.0.0/0
e) ip prefix-list E permit 0.0.0.0/1 le 24
a) All host routes
b) Any subnet in class B address space
c) All routes
d) Just the default route
e) Any prefix in class A address space covering at least 256 addresses
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-14

The figure above contains some commonly used prefix-list examples.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) When you are defining prefix-lists, what are two reasons to use sequence numbers? (Choose two.)

A) to reference the associated ACL for the prefix-list entry

B) to provide a means of linking an AS-path filter-list to the prefix-list

C) to provide an execution order for prefix-list entries

D) to provide a means of inserting or deleting list entries

Q2) When you are defining a prefix-list using the *ge-value* or *le-value*, what is true regarding the *len* parameter?

A) *ge-value* is greater than or equal to *le-value*

B) *le-value* is less than or equal to *len*

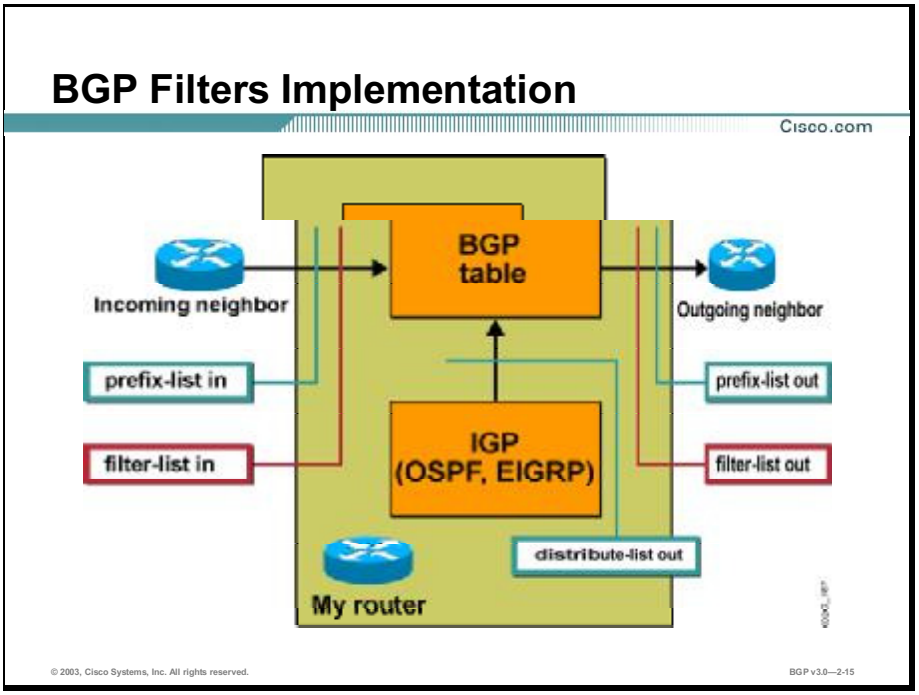
C) *len* is less than *ge-value* is less than or equal to *le-value* is less than or equal to 32

D) 8 is less than or equal to *len* is less than or equal to *ge-value* is less than or equal to 32

Copyright © 2003, Cisco Systems, Inc.

BGP Filters Implementation

This topic describes where you can implement prefix-lists in a BGP network.



You can optionally apply filter-lists and prefix-lists on either incoming or outgoing neighbors in any combination. Both the incoming prefix-list and the incoming filter-list must permit the routes received from a neighbor before being accepted into the BGP table. Outgoing routes must pass both the outgoing filter-list and the outgoing prefix-list before being transmitted to the neighbor.

When a router is configured to redistribute routing information from an Interior Gateway Protocol (IGP) into BGP, the routes must successfully pass any prefix-list or access-list applied to the redistribution before a route is injected into the BGP table.

Practice

Q1) How are prefix-lists and the AS-path filters combined?

A) They cannot be combined, because the router can use either prefix-lists or AS-path filters but not both simultaneously.

B) The prefix-list and AS-path lists use the same filter-list.

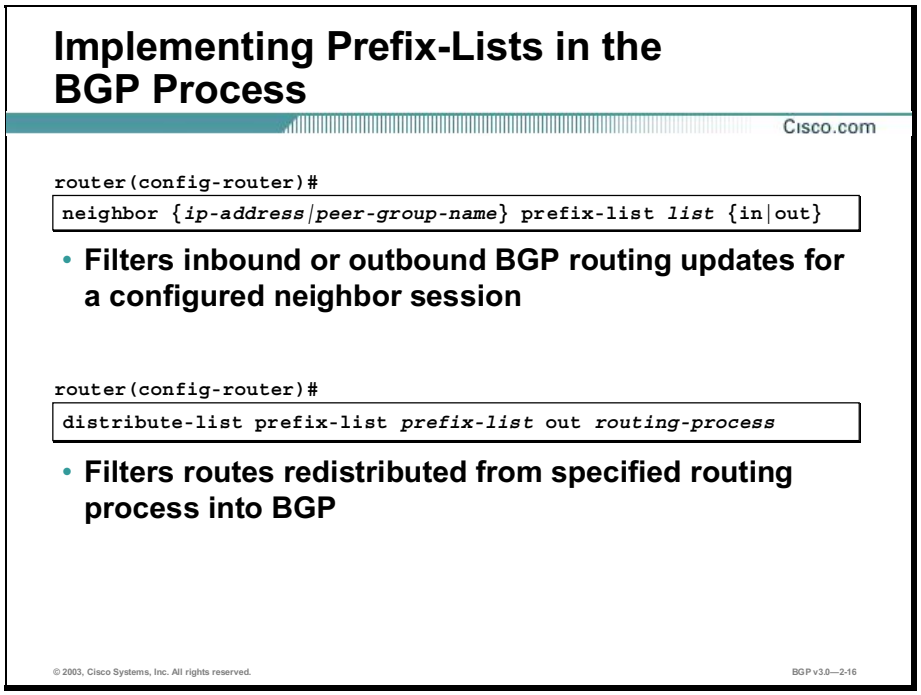
C) The router checks both lists for the first match and then stops processing.

D) All updates are matched against the prefix-list and the AS-path filter before being sent or accepted.

 Copyright © 2003, Cisco Systems, Inc.

Implementing Prefix-Lists in the BGP Process

This topic identifies the Cisco IOS commands required to apply prefix-lists for filtering of inbound or outbound updates.


Implementing Prefix-Lists in the BGP Process
Cisco.com
router(config-router)#
neighbor {ip-address|peer-group-name} prefix-list list {in|out}
• Filters inbound or outbound BGP routing updates for a configured neighbor session
router(config-router)#
distribute-list prefix-list prefix-list out routing-process
• Filters routes redistributed from specified routing process into BGP
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-16

You can use prefix-lists to filter incoming or outgoing BGP updates to neighbors. You can also use them to filter routes being redistributed into the BGP process from other routing protocols.

neighbor prefix-list

To distribute BGP neighbor information as specified in a prefix-list, use the **neighbor prefix-list** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **prefix-list** *prefix-listname* {**in** | **out**}

To remove an entry, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **prefix-list** *prefix-listname* {**in** |**out**}

Syntax Description

ip-address	IP address of neighbor.
peer-group-name	Name of a BGP peer group.
prefix-listname	Name of a prefix-list.

in	Access list is applied to incoming advertisements to that neighbor.
out	Access list is applied to outgoing advertisements from that neighbor.

Note A BGP peer group is a group of BGP neighbors with the same update policies. Route-maps, distribute-lists, filter-lists, and so on usually set update policies. Instead of defining the same policies for each separate neighbor, a peer group name is configured on the router, and these policies are assigned to the peer group. BGP peer groups are discussed in the Advanced BGP module.

distribute-list out

To suppress networks from being advertised in updates, use the **distribute-list out** router configuration command.

distribute-list {*access-list-number* | *name* | **prefix-list** *prefix-listname*} **out** [*interface-name* | *routing-process* | *autonomous-system-number*]

To disable this function, use the **no** form of this command.

no distribute-list {*access-list-number* | *name* | **prefix-list** *prefix-listname*} **out** [*interface-name* | *routing-process* | *autonomous-system-number*]

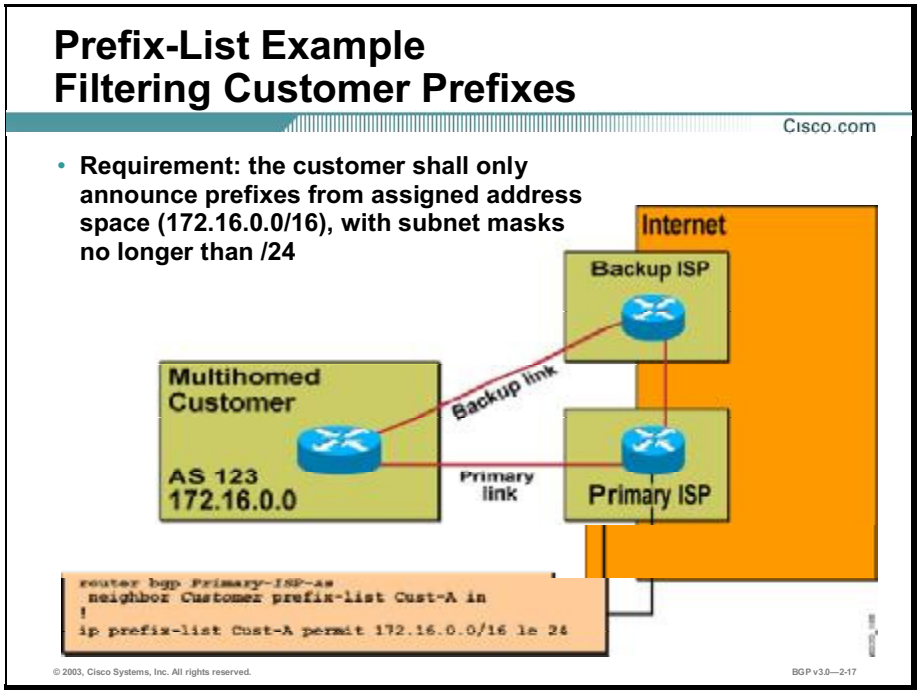
Syntax Description

access-list-number | *name*	Standard IP access-list number or name. The list defines which networks are to be received and which are to be suppressed in routing updates.
prefix-listname	Name of a prefix-list. The list defines which networks are to be received and which are to be suppressed in routing updates, based upon matching of the network prefix to the prefixes in the list.
out	Applies the access-list to outgoing routing updates.
interface-name	(Optional) Name of a particular interface.
routing-process	(Optional) Name of a particular routing process, or the keyword **static** or **connected**.
autonomous-system-number	(Optional) Autonomous system (AS) number

Note Although you can use the **neighbor prefix-list** router configuration command as an alternative to the **neighbor distribute-list** command, do not use both the **neighbor prefix-list** and **neighbor distribute-list** command filtering to the same neighbor in any given direction. These two commands are mutually exclusive, and only one command (**neighbor prefix-list** or **neighbor distribute-list**) can be applied for each inbound or outbound direction.

 Copyright © 2003, Cisco Systems, Inc.

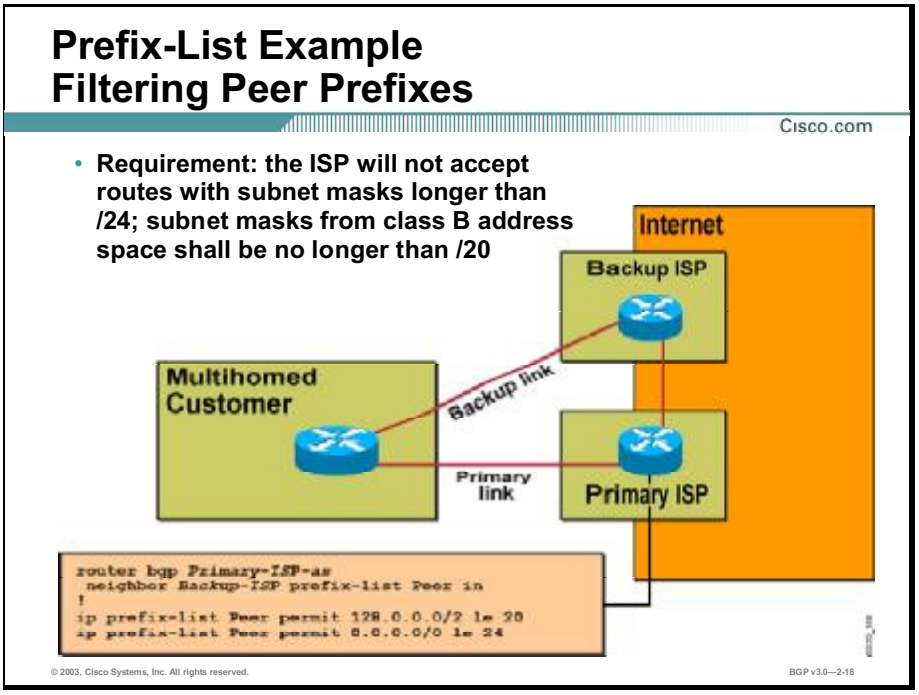
Example



In this example, a multihomed customer has been assigned the address space 172.16.0.0/16. The customer may subnet this address space but may not announce smaller subnets than a subnet mask of 255.255.255.0. Larger subnets are accepted. If the customer has subnetted the network into smaller subnets, it must summarize the routing information about those subnets into at least /24 prefixes before announcing them.

The primary ISP implements a prefix-list named “Cust-A” to perform the filtering of incoming information from the multihomed customer. The prefix-list permits all routes received from the customer that have 172.16 in the first 16 bits and have a subnet mask of 24 bits or less. Any other routes from the customer are denied and silently ignored.

Example


Prefix-List Example
Filtering Peer Prefixes
Cisco.com
• Requirement: the ISP will not accept routes with subnet masks longer than /24; subnet masks from class B address space shall be no longer than /20
Internet
Backup ISP
Multihomed Customer
Backup link
Primary link
Primary ISP
router bgp Primary-ISP-as
neighbor Backup-ISP prefix-list Peer in
!
ip prefix-list Peer permit 128.0.0.0/2 le 20
ip prefix-list Peer permit 0.0.0.0/0 le 24
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-18

The primary ISP will not accept any route from the customer that indicates a subnet smaller than a 255.255.255.0 subnet mask. The class B network, however, must not be subnetted into smaller subnets than a 255.255.240.0 subnet mask.

The primary ISP implements this route by using a prefix-list named “Peer.” The first line in the prefix-list checks if it is a class B network. Remember that a class B address always has the binary sequence 10 as the first two bits in the first byte. The second line matches any prefix.

When the primary ISP receives a route from the customer, it compares the route with both lines. If the route is a class B network, both lines match. Testing will continue by checking the subnet mask. An upper bound is explicitly indicated, giving a maximum prefix length of 20 bits.

If the received route is not a class B network, only the second line matches. In this case, the subnet mask length must be greater or equal to 0 and less than or equal to 24, giving a route less explicit than a /24 prefix.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) How can you apply the same prefix-list to multiple BGP neighbors on a router?

A) by configuring a **neighbor prefix-list** statement for each BGP peer

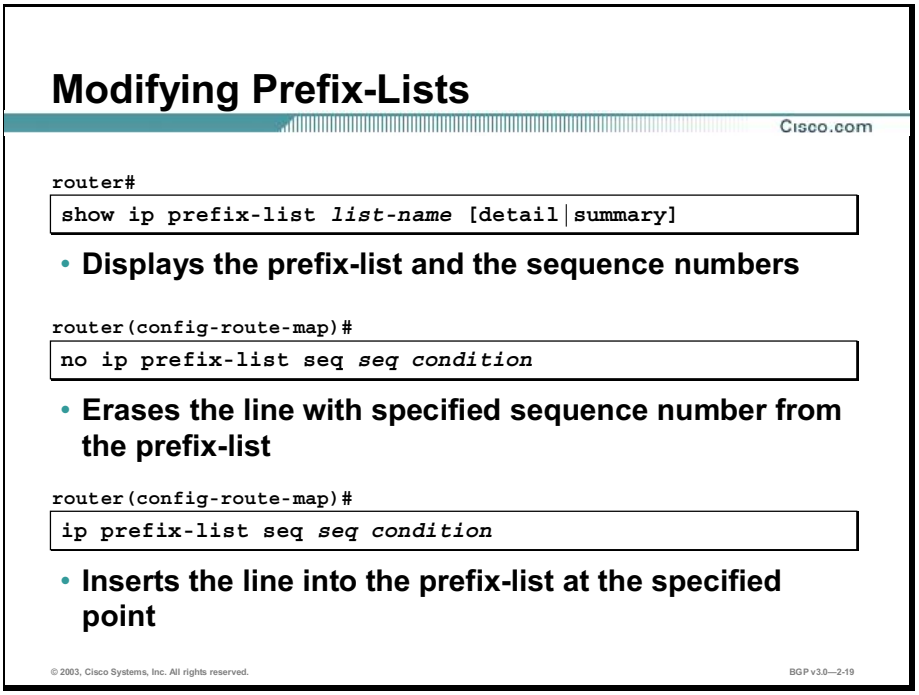
B) by configuring a **neighbor distribute-list** statement for each neighbor

C) by using the BGP *peer-group* option with the **neighbor** statement

D) by configuring the prefix-list as a global filter under the BGP routing process

Modifying Prefix-Lists

This topic identifies the Cisco IOS commands required to modify configured prefix-list filters.


Modifying Prefix-Lists
Cisco.com
router#
show ip prefix-list list-name [detail|summary]
• Displays the prefix-list and the sequence numbers
router(config-route-map)#
no ip prefix-list seq seq condition
• Erases the line with specified sequence number from the prefix-list
router(config-route-map)#
ip prefix-list seq seq condition
• Inserts the line into the prefix-list at the specified point
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-19

Lines in a prefix-list are assigned sequence numbers. These assignments significantly increase the manageability of the list. They provide the opportunity to remove a specific line, and, if spacing between the sequence numbers allows, they provide the ability to insert a line between two existing lines.

To display a currently configured prefix-list and its sequence numbers, use the **show ip prefix-list** command with the **detail** keyword.

You can specify sequence values for prefix-list entries in any increments you want (the automatically generated numbers are incremented in units of 5). If you specify the sequence values in increments of 1, you will not be able to insert additional entries into the prefix-list. If you choose very large increments, you could run out of sequence values.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) If sequence number 2 is inserted in the first line of a prefix-list, but no other sequence numbers are entered on subsequent lines, what will the sequence numbers be for the next three statements?

A) This will generate an error, because all remaining sequence numbers must be manually entered.

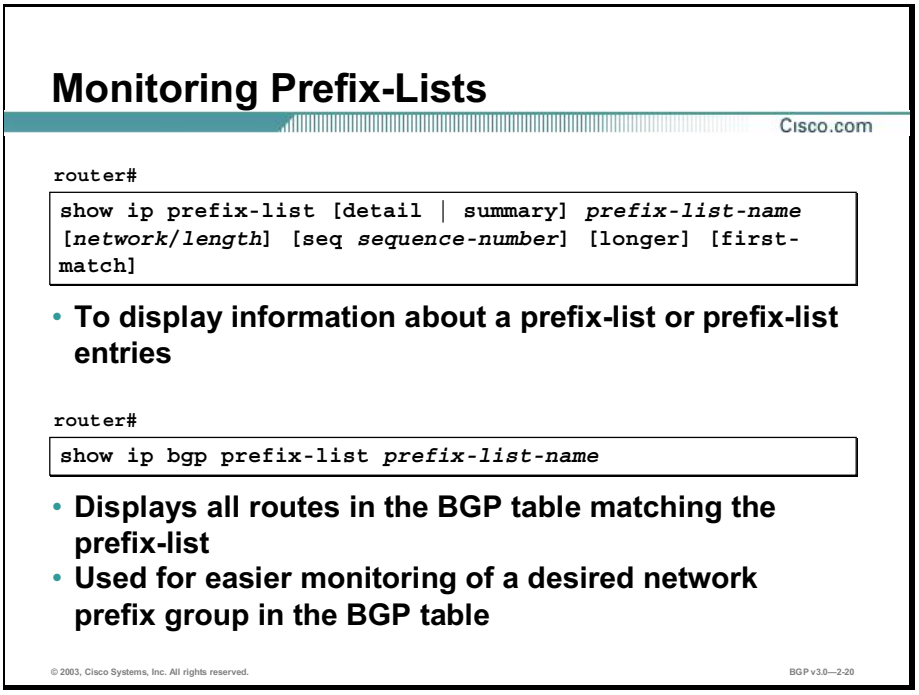
B) 3, 4, 5

C) 7, 12, 17

D) 5, 10, 15

Monitoring Prefix-Lists

This topic lists the Cisco IOS commands required to monitor the operation of configured prefix-list filters.


Monitoring Prefix-Lists
Cisco.com
router#
show ip prefix-list [detail | summary] prefix-list-name [network/length] [seq sequence-number] [longer] [first-match]
• To display information about a prefix-list or prefix-list entries
router#
show ip bgp prefix-list prefix-list-name
• Displays all routes in the BGP table matching the prefix-list
• Used for easier monitoring of a desired network prefix group in the BGP table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-20

show ip prefix-list

To display information about a prefix-list or prefix-list entries, use the **show ip prefix-list** EXEC command.

show ip prefix-list [**detail** |**summary**] *name* [*network/len*] [**seq** *seq-num*] [**longer**] [**first-match**]

Syntax Description

detail | **summary**	(Optional) Displays detailed or summarized information about all prefix-lists
name	(Optional) The name of a specific prefix-list
network/len	(Optional) The network number and length (in bits) of the network mask
seq	(Optional) Applies the sequence number to the prefix-list entry
seq-num	The sequence number of the prefix-list entry
longer	Displays all entries of a prefix-list that are more specific than the given *network/len*
first-match	Displays the entry of a prefix-list that matches the given *network/len*

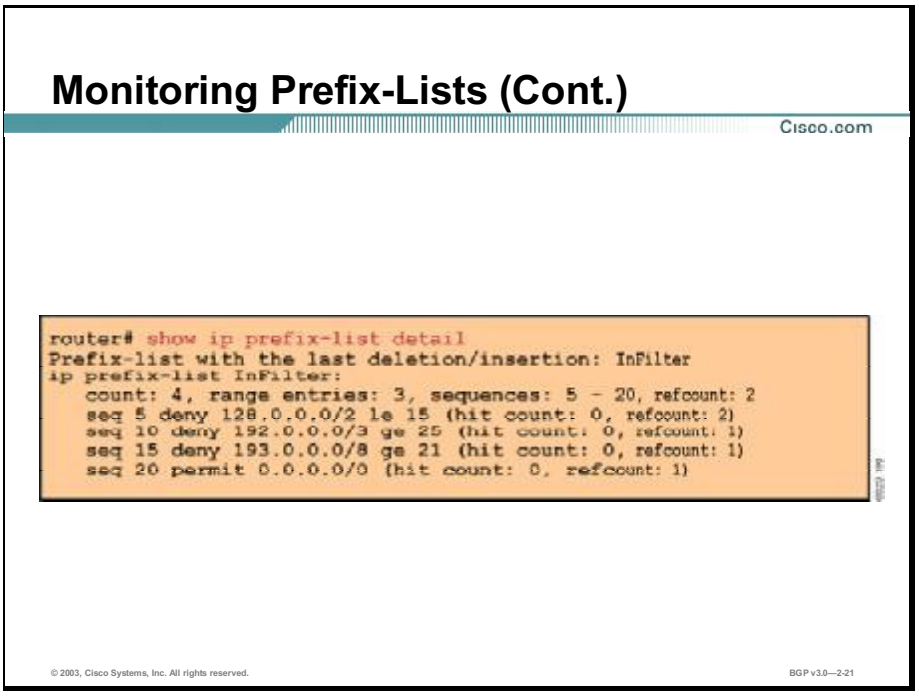
 Copyright © 2003, Cisco Systems, Inc.

The **show ip bgp prefix-list** command displays selected routes from a BGP routing table based on the contents of a prefix-list. Use it for selective filtering of BGP table output on Cisco IOS devices on the basis of network prefix groups.

To perform prefix-list-based BGP table filtering, follow these steps:

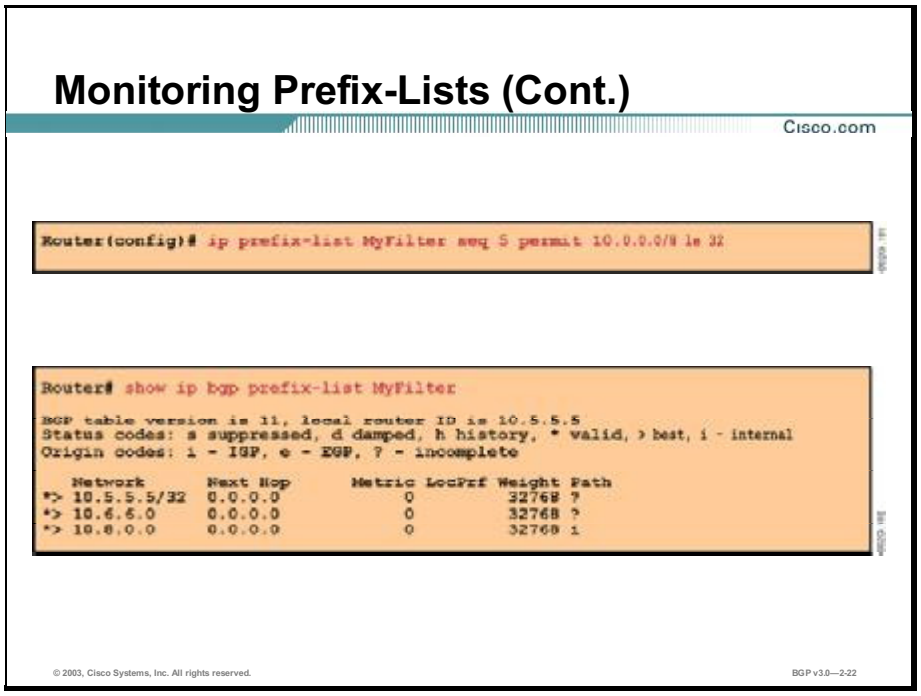
- Configure a prefix-list permitting ranges of networks meant to be displayed in the BGP table output
- Include a reference to a configured prefix-list in the **show ip bgp prefix-list** command

Note The support for prefix-list BGP table filtering was added in 12.2(11)T and 12.0(14)ST images of Cisco IOS software.


Monitoring Prefix-Lists (Cont.)
Cisco.com
router# show ip prefix-list detail
Prefix-list with the last deletion/insertion: InFilter
ip prefix-list InFilter:
count: 4, range entries: 3, sequences: 5 - 20, refcount: 2
seq 5 deny 128.0.0.0/2 le 15 (hit count: 0, refcount: 2)
seq 10 deny 192.0.0.0/3 ge 25 (hit count: 0, refcount: 1)
seq 15 deny 193.0.0.0/8 ge 21 (hit count: 0, refcount: 1)
seq 20 permit 0.0.0.0/0 (hit count: 0, refcount: 1)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-21

In this example, the **show ip prefix-list** command has been issued with the **detail** keyword. The output of the command displays detailed information about configured prefix-lists, including sequence numbers, the prefix-list entries, and the number of times that each entry has been matched by a corresponding prefix.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Prefix-Lists (Cont.)
Cisco.com
Router(config)# ip prefix-list MyFilter seq 5 permit 10.0.0.0/8 le 32
Router# show ip bgp prefix-list MyFilter
BGP table version is 11, local router ID is 10.5.5.5
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 10.5.5.5/32 0.0.0.0 0 32768 ?
*> 10.6.6.0 0.0.0.0 0 32768 ?
*> 10.8.0.0 0.0.0.0 0 32768 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-22

This example shows a simple prefix-list–based filtering of the BGP table. The prefix-list filter will permit all networks with the first octet equal to 10 and any length of a subnet mask (**le 32**). In the **show ip bgp prefix-list** output, only the networks permitted by the prefix-list filter are displayed.

Practice

Q1) How can you use the **show ip prefix-list** command to display the prefix-list entry matching a specific prefix and length?

A) This is not a feature of the **show ip prefix-list** command.

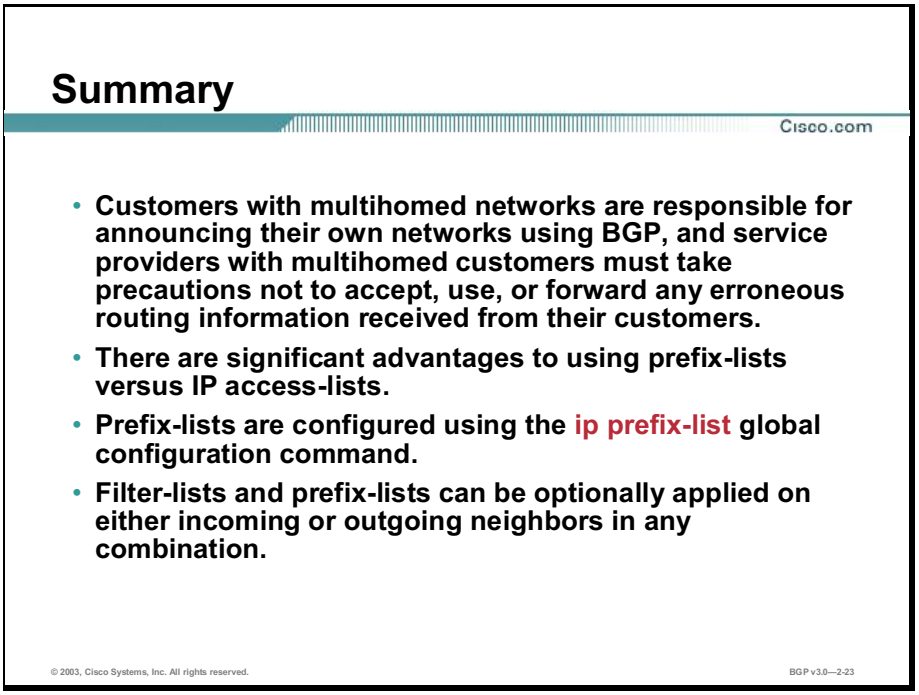
B) By specifying the **detail** keyword.

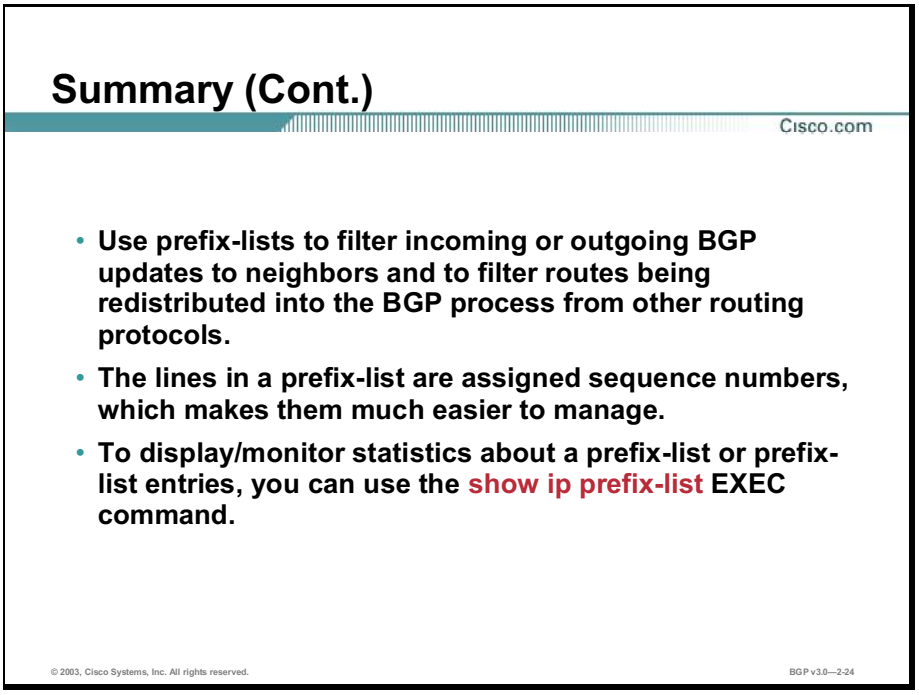
C) With the longer keyword to display all matches except those with more specific entries.

D) By specifying the **first-match** keyword.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Customers with multihomed networks are responsible for announcing their own networks using BGP, and service providers with multihomed customers must take precautions not to accept, use, or forward any erroneous routing information received from their customers.
• There are significant advantages to using prefix-lists versus IP access-lists.
• Prefix-lists are configured using the ip prefix-list global configuration command.
• Filter-lists and prefix-lists can be optionally applied on either incoming or outgoing neighbors in any combination.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-23


Summary (Cont.)
Cisco.com
• Use prefix-lists to filter incoming or outgoing BGP updates to neighbors and to filter routes being redistributed into the BGP process from other routing protocols.
• The lines in a prefix-list are assigned sequence numbers, which makes them much easier to manage.
• To display/monitor statistics about a prefix-list or prefix-list entries, you can use the show ip prefix-list EXEC command.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-24

 Copyright © 2003, Cisco Systems, Inc.

Next Steps

After completing this lesson, go to:

- Outbound Route Filtering lesson

References

For additional information, refer to these resources:

- For more information on prefix-list filters, refer to “Configuring BGP” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/np1_c/1cprt1/1cbgp.htm#xtocid15

Laboratory Exercise: Prefix-List Filters

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using prefix-list filters where you must support connections to multiple ISPs.

After completing this exercise, you will be able to:

- Configure filtering of inbound or outbound updates with prefix-lists
- Modify configured prefix-list filters
- Monitor the operations of configured prefix-list filters

 Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
ip prefix-list *name* [**seq** *n*] {**permit**|**deny**} **prefix/length** [**ge** *length* | **le** *length*]	Enter a prefix list.
neighbor *ip-address* **prefix-list** *prefix-list* **in**	Filter incoming prefixes.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- The memory consumption on WGxR1 has increased due to the large BGP tables received from your service providers. You realize that the service providers announce a large number of very small prefixes to you. Some of the announced prefixes are also in the private IP (RFC 1918) address space.
- In this exercise, you will minimize the size of the BGP table on your router with inbound filters implemented with prefix-lists. You will also filter any prefixes in the private IP address space that you might receive from your service providers.
- Your analysis shows that you must perform the following tasks to minimize the size of the BGP table on router WGxR1:
  - Do not accept any subnets of class B networks from router “Cheap.”
  - Do not accept class C networks or its subnets from router “Good.”
  - Do not accept any private networks (RFC 1918) from either neighbor.
  - Use prefix-lists to perform these tasks.
- Figure 1 shows the connectivity that is established between your AS and the two service providers “Good” and “Cheap.”

Copyright © 2003, Cisco Systems, Inc.

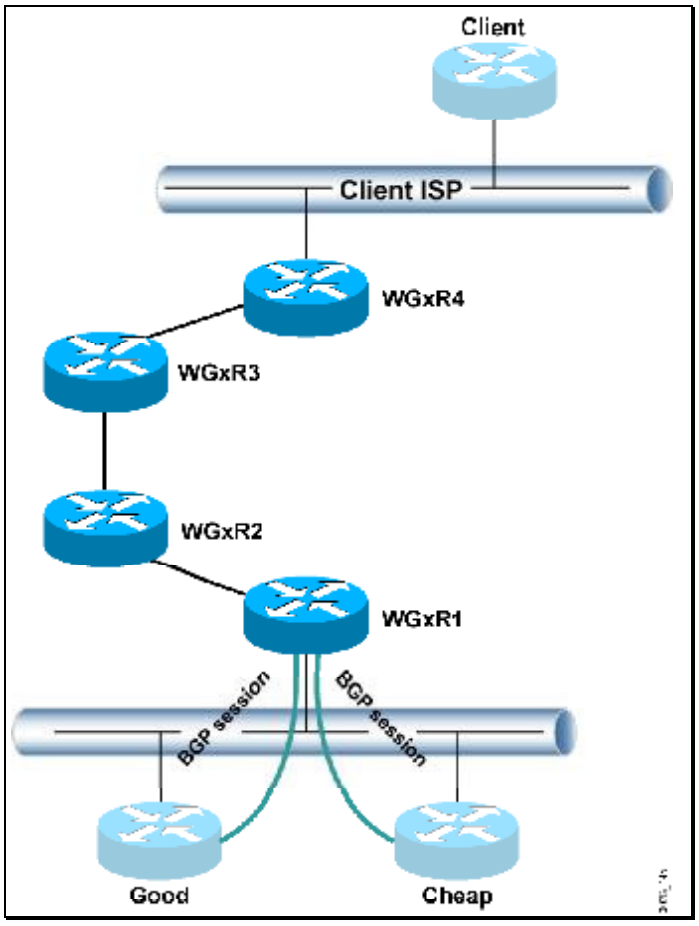


Figure 1: Connections to two different service providers

Exercise Procedure

Complete these steps:

Step 1 Verify that your current BGP table looks somewhat like the following:

```
wg1r1#sh ip bgp
BGP table version is 44, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.0.0.0         192.168.20.233           0             0 20 i
*>                  192.168.20.233                       100 22 20 i
*  128.20.0.0       192.168.20.20            0             0 20 i
*>                  192.168.20.22            0           100 22 i
*  128.20.12.0/24   192.168.20.20            0             0 20 i
*>                  192.168.20.20                        100 22 20 i
*  128.22.0.0       192.168.20.20            0             0 20 i
*>                  192.168.20.22            0           100 22 i
*  128.22.12.0/24   192.168.20.22                          0 20 22 i
*>                  192.168.20.22            0           100 22 i
*  128.26.0.0       192.168.20.20            0             0 20 42 26
i
*>                  192.168.20.22            0           100 22 26 i
*  128.37.0.0       192.168.20.20            0             0 20 42 37
i
```

 Copyright © 2003, Cisco Systems, Inc.

```
*>                  192.168.20.22           0           100 22 26 42
37 i
*  128.42.0.0       192.168.20.20           0             0 20 42 i
*>                  192.168.20.22           0           100 22 26 42
i
*  128.51.0.0       192.168.20.20           0             0 20 42 26
51 i
*>                  192.168.20.22           0           100 22 26 51
i
*> 128.213.0.0      192.168.20.20           0             0 20 213 i
*> 128.214.0.0      192.168.20.22           0           100 22 214 i
*  192.20.11.0      192.168.20.20           0             0 20 i
*>                  192.168.20.22           0           100 22 i
*  192.20.12.0/30   192.168.20.20           0             0 20 i
*>                  192.168.20.20                       100 22 20 i
*  192.22.11.0      192.168.20.20           0             0 20 i
*>                  192.168.20.22           0           100 22 i
*  192.22.12.0/30   192.168.20.22                         0 20 22 i
*>                  192.168.20.22           0           100 22 i
*  192.26.11.0      192.168.20.20           0             0 20 42 26
i
*>                  192.168.20.22           0           100 22 26 i
*  192.37.11.0      192.168.20.20           0             0 20 42 37
i
*>                  192.168.20.22           0           100 22 26 42
37 i
*  192.42.11.0      192.168.20.20           0             0 20 42 i
*>                  192.168.20.22           0           100 22 26 42
i
*  192.51.11.0      192.168.20.20           0             0 20 42 26
51 i
*>                  192.168.20.22           0           100 22 26 51
i
*> 192.168.1.0      0.0.0.0                 0         32768 i
*> 192.213.11.0     192.168.20.20           0             0 20 213 i
*> 192.214.11.0     192.168.20.22           0           100 22 214 i
*> 197.1.0.0/21     0.0.0.0                           32768 i
*> 197.1.0.0/16     0.0.0.0                           32768 i
s> 197.1.1.0        0.0.0.0                 0         32768 i
s> 197.1.2.0        0.0.0.0                 0         32768 i
s> 197.1.3.0        0.0.0.0                 0         32768 i
s> 197.1.4.0        0.0.0.0                 0         32768 i
s> 197.1.5.0        0.0.0.0                 0         32768 i
s> 197.1.6.0        0.0.0.0                 0         32768 i
s> 197.1.7.0        0.0.0.0                 0         32768 i
s> 197.1.8.0        0.0.0.0                 0         32768 i
*> 197.1.8.0/22     0.0.0.0                           32768 i
*  200.20.0.0/16    192.168.20.20           0             0 20 i
*>                  192.168.20.20                       100 22 20 i
*  200.22.0.0/16    192.168.20.22                         0 20 22 i
*>                  192.168.20.22           0           100 22 i
```

Configuring prefix-lists:

Step 2 Based on your analysis, create prefix-lists to filter incoming updates from routers “Good” and “Cheap.”

Applying prefix-lists:

Step 3 Apply the prefix-lists that you have created to updates coming from routers “Good” and “Cheap.”

Exercise Verification

You have completed this exercise when you attain these results:

- Inspect your BGP table and check the proper operation of your prefix-list filters.

```
wg1r1#sh ip bgp reg ^20_
BGP table version is 50, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  128.20.0.0       192.168.20.20            0             0 20 i
*> 128.20.12.0/24   192.168.20.20            0             0 20 i
*  128.22.0.0       192.168.20.20            0             0 20 i
*> 128.22.12.0/24   192.168.20.22                          0 20 22 i
*  128.26.0.0       192.168.20.20            0             0 20 42 26
i
*  128.37.0.0       192.168.20.20            0             0 20 42 37
i
*  128.42.0.0       192.168.20.20            0             0 20 42 i
*  128.51.0.0       192.168.20.20            0             0 20 42 26
51 i
*> 128.213.0.0      192.168.20.20            0             0 20 213 i
*  200.20.0.0/16    192.168.20.20            0             0 20 i
*  200.22.0.0/16    192.168.20.22                          0 20 22 i
wg1r1#
```

Full BGP table:

```
wg1r1#sh ip bgp
BGP table version is 50, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 128.20.0.0       192.168.20.22            0           100 22 i
*                   192.168.20.20            0             0 20 i
*> 128.20.12.0/24   192.168.20.20            0             0 20 i
*> 128.22.0.0       192.168.20.22            0           100 22 i
*                   192.168.20.20            0             0 20 i
*> 128.22.12.0/24   192.168.20.22                          0 20 22 i
*> 128.26.0.0       192.168.20.22            0           100 22 26 i
*                   192.168.20.20            0             0 20 42 26
i
*> 128.37.0.0       192.168.20.22            0           100 22 26 42
37 i
*                   192.168.20.20            0             0 20 42 37
i
*> 128.42.0.0       192.168.20.22            0           100 22 26 42
i
*                   192.168.20.20            0             0 20 42 i
*> 128.51.0.0       192.168.20.22            0           100 22 26 51
i
*                   192.168.20.20            0             0 20 42 26
51 i
*> 128.213.0.0      192.168.20.20            0             0 20 213 i
*> 128.214.0.0      192.168.20.22            0           100 22 214 i
*> 192.20.11.0      192.168.20.22            0           100 22 i
*> 192.20.12.0/30   192.168.20.20                        100 22 20 i
*> 192.22.11.0      192.168.20.22            0           100 22 i
*> 192.22.12.0/30   192.168.20.22            0           100 22 i
*> 192.26.11.0      192.168.20.22            0           100 22 26 i
```

 Copyright © 2003, Cisco Systems, Inc.

```
*> 192.37.11.0      192.168.20.22            0           100 22 26 42
37 i
*> 192.42.11.0      192.168.20.22            0           100 22 26 42
i
*> 192.51.11.0      192.168.20.22            0           100 22 26 51
i
*> 192.168.1.0      0.0.0.0                  0         32768 i
*> 192.214.11.0     192.168.20.22            0           100 22 214 i
*> 197.1.0.0/21     0.0.0.0                            32768 i
*> 197.1.0.0/16     0.0.0.0                            32768 i
s> 197.1.1.0        0.0.0.0                  0         32768 i
s> 197.1.2.0        0.0.0.0                  0         32768 i
s> 197.1.3.0        0.0.0.0                  0         32768 i
s> 197.1.4.0        0.0.0.0                  0         32768 i
s> 197.1.5.0        0.0.0.0                  0         32768 i
s> 197.1.6.0        0.0.0.0                  0         32768 i
s> 197.1.7.0        0.0.0.0                  0         32768 i
s> 197.1.8.0        0.0.0.0                  0         32768 i
*> 197.1.8.0/22     0.0.0.0                            32768 i
*> 200.20.0.0/16    192.168.20.20                        100 22 20 i
*                   192.168.20.20            0             0 20 i
*> 200.22.0.0/16    192.168.20.22            0           100 22 i
*                   192.168.20.22                          0 20 22 i
```

Answer these questions:

Q1) Are neighboring workgroups still reachable?

Q2) Why does router WGxR1 still accept class B networks from router “Cheap”?

 Copyright © 2003, Cisco Systems, Inc.

Outbound Route Filtering

Overview

Outbound route filtering is a prefix-based feature that minimizes the number of Border Gateway Protocol (BGP) updates that are sent between peer routers. An outbound route filter (ORF) aids in the reduction of resources required for generating and processing routing updates by filtering out unwanted routing updates at the source. One common use of BGP outbound route filtering is to reduce the amount of processing required on a router that is not accepting full routes from a service provider network to which it is connected.

This lesson discusses the function of outbound route filtering in a BGP network. The format and function of ORF messages are discussed as well as the commands that enable ORF negotiations and the activation of an ORF prefix-list. The commands used to trigger a route refresh will be detailed and finally a discussion on how to monitor the operations of a configured ORF in a BGP network.

Importance

An ORF is an additional mechanism used to minimize the number of updates being requested from a neighbor, which reduces link bandwidth consumption and CPU use when a router requests a route refresh. ORF also allows filtering of information that external networks should not receive (such as RFC 1918 information). Understanding how to monitor ORF capabilities is also important because a BGP neighbor supporting specific ORF capabilities will report those capabilities to a monitoring neighbor and can then send a filter of the supported type to the neighbor.

Objectives

Upon completing this lesson, you will be able to:

- Identify the function of outbound route filtering in a BGP network
- Describe the format and function of an ORF message
- Identify the Cisco IOS® commands required to enable ORF negotiations and activate an ORF prefix-list
- Identify the Cisco IOS commands used to trigger a route refresh
- Identify the Cisco IOS commands required to monitor the operation of a configured ORF

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

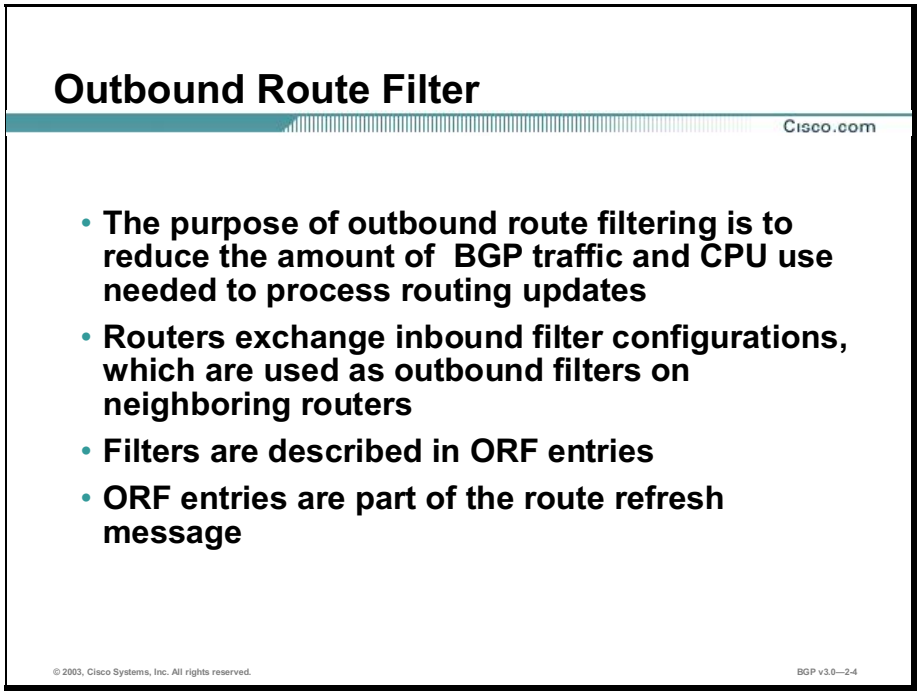
This lesson includes these topics:

- Overview
- Outbound Route Filter
- Outbound Route Filter Message
- Configuring Outbound Route Filtering
- Using Outbound Route Filtering
- Monitoring Outbound Route Filtering
- Summary
- Assessment (Quiz): Outbound Route Filtering

 Copyright © 2003, Cisco Systems, Inc.

Outbound Route Filter

This topic identifies the function of outbound route filtering in a BGP network.


Outbound Route Filter
Cisco.com
• The purpose of outbound route filtering is to reduce the amount of BGP traffic and CPU use needed to process routing updates
• Routers exchange inbound filter configurations, which are used as outbound filters on neighboring routers
• Filters are described in ORF entries
• ORF entries are part of the route refresh message
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-4

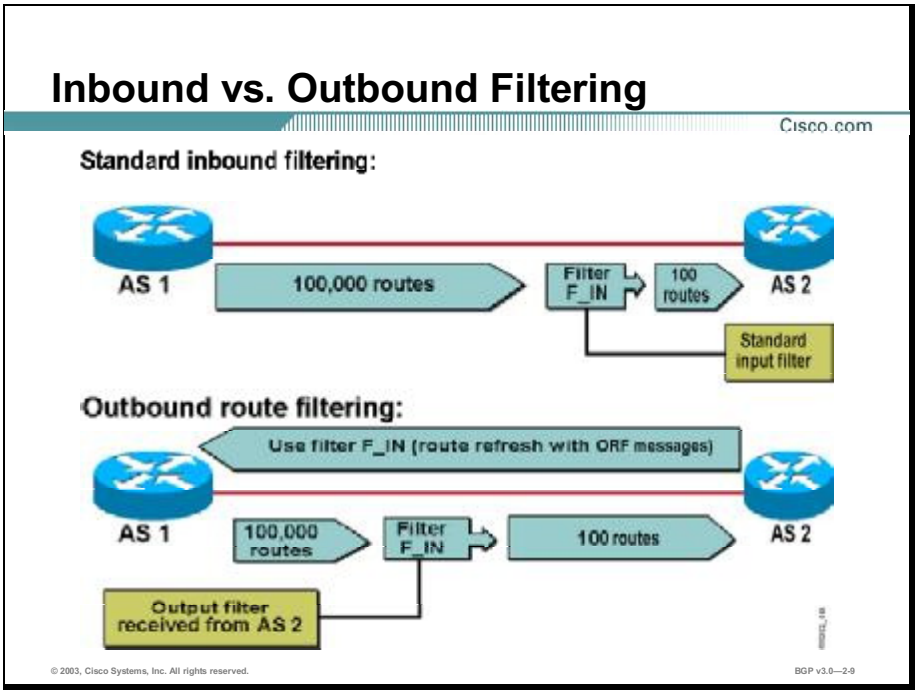
Outbound route filtering is a prefix-based BGP feature that is enabled through the advertisement of ORF capabilities to peer routers. The advertisement of the ORF capability indicates that a BGP-speaking router will accept a prefix-list from a neighbor and apply the prefix-list to locally configured ORFs (if any exist). When this capability is enabled, the BGP speaker can install an inbound prefix-list filter to the remote peer as an outbound filter, which reduces unwanted routing updates.

The standard route refresh message contains the address family information for which the refresh is needed. Outbound route filtering is an additional mechanism used to minimize the number of updates being requested from a neighbor.

This mechanism reduces link bandwidth consumption and CPU use when a router requests a route refresh. Filters that should be used by routers with the route refresh are described in ORF entries that are part of the route refresh message.

You can configure the ORF feature with send, receive, or send and receive capabilities. The local peer advertises the ORF capability in send mode. The remote peer receives the ORF capability in receive mode and applies the filter as outbound policy. The local and remote peers exchange updates to maintain the ORF for each router. Peer routers exchange updates depending on the ORF prefix-list capability that is advertised. The remote peer starts sending updates to the local peer after it receives a route refresh request or an ORF prefix-list with an immediate status.

Example



Outbound route filtering can limit the number of unwanted routing updates, which will reduce the amount of resources required for routing update generation and processing. This feature also reduces the amount of resources required to receive and discard routes that would otherwise be filtered out by the receiving router if the ORF feature was not available.

The example shows two scenarios:

Step 1 The upper example shows that 100,000 routes are sent to a neighbor, and the input filter permits only 100 of these routes.

Step 2 The second example shows how a route refresh with a filter is sent to the neighbor. The neighbor then uses the filter before sending the updates. This way only 100 updates are sent.

Practice

Q1) What are two benefits of using the outbound route filtering feature? (Choose two.)

A) It minimizes the amount of BGP routing traffic transferred over the directly connected link.

B) It defines inbound filters used on remote peer routers.

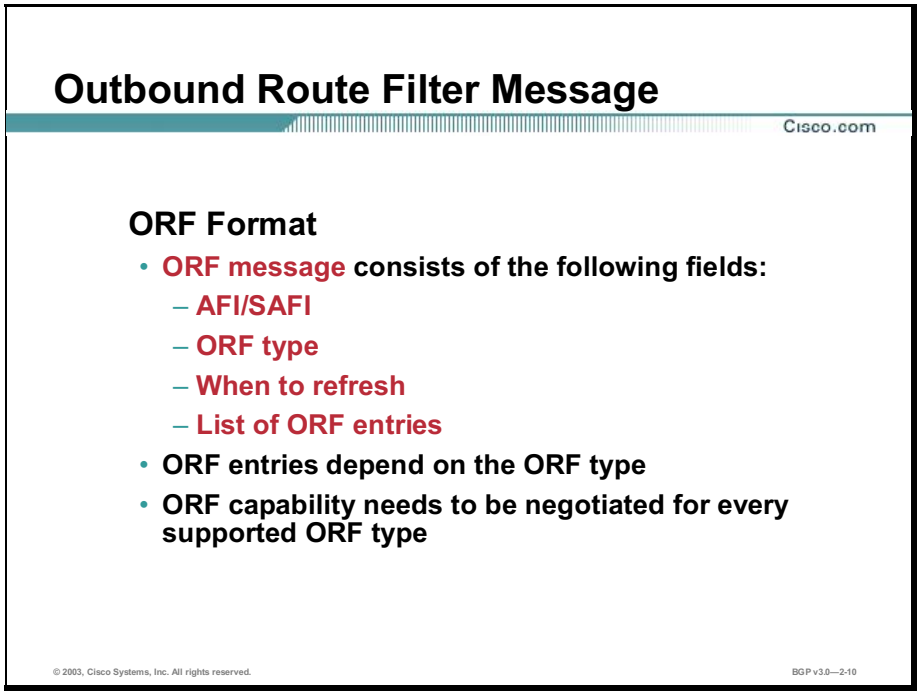
C) It reduces the amount of CPU consumed by a router.

D) You can use ORF to implement nondisruptive changes in BGP policy.

 Copyright © 2003, Cisco Systems, Inc.

Outbound Route Filter Message

This topic describes the format and function of the ORF message.

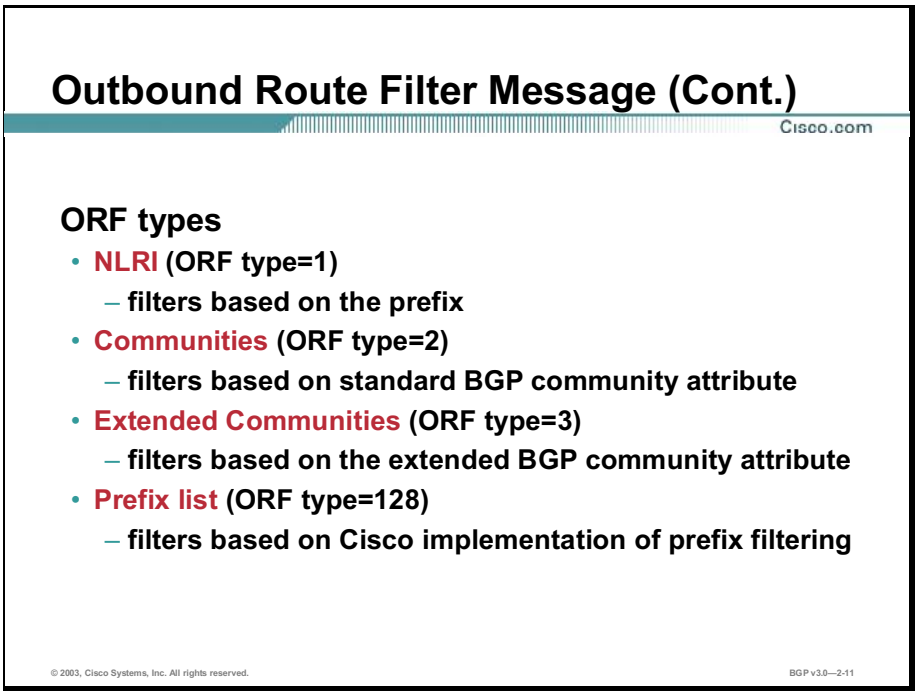

Outbound Route Filter Message
Cisco.com
ORF Format
• ORF message consists of the following fields:
– AFI/SAFI
– ORF type
– When to refresh
– List of ORF entries
• ORF entries depend on the ORF type
• ORF capability needs to be negotiated for every supported ORF type
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-10

An ORF message contains the following information:

- Address Family Information (AFI) and Subsequent Address Family Information (SAFI) for which the filter should be used
- ORF type, which identifies the type of the filter
- When to refresh (immediate or deferred refresh)
- List of ORF entries where the actual filter is defined

You can use the AFI/SAFI component of the ORF message to provide a coarse level of granular control by limiting the ORF to only the routes whose Network Layer Reachability Information (NLRI) matches the configured AFI/SAFI component.

ORF capability has to be negotiated by the router for each ORF type supported in the ORF message.

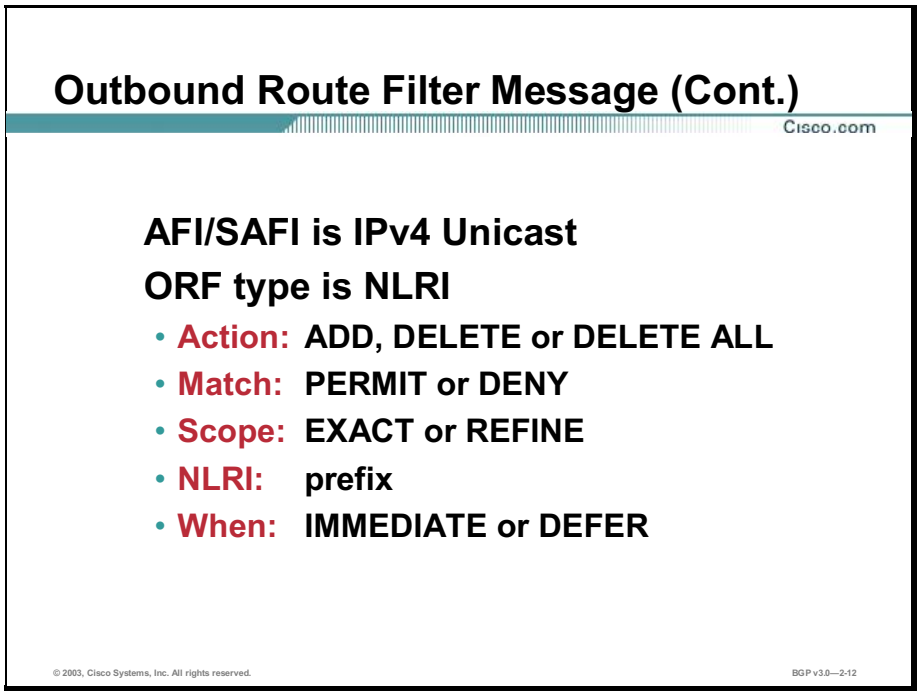

Outbound Route Filter Message (Cont.)
Cisco.com
ORF types
• NLRI (ORF type=1)
– filters based on the prefix
• Communities (ORF type=2)
– filters based on standard BGP community attribute
• Extended Communities (ORF type=3)
– filters based on the extended BGP community attribute
• Prefix list (ORF type=128)
– filters based on Cisco implementation of prefix filtering
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-11

The value contained in the ORF type determines the content contained in the ORF message.

Currently, ORF type 0 is reserved, ORF types 1 to 127 are assigned by the Internet Assigned Numbers Authority (IANA), and ORF types 128 to 255 are vendor-specific (and not assigned by the IANA). Commonly used ORF types are as follows:

- ORF type 1 is used to filter based on NLRI.
- ORF type 2 is used to filter based on standard BGP community attributes.
- ORF type 3 is used to filter based on extended BGP community attributes.
- ORF type 128 is used to filter based on Cisco proprietary implementation of prefix filtering (prefix-lists).

 Copyright © 2003, Cisco Systems, Inc.


Outbound Route Filter Message (Cont.)
Cisco.com
AFI/SAFI is IPv4 Unicast
ORF type is NLRI
• Action: ADD, DELETE or DELETE ALL
• Match: PERMIT or DENY
• Scope: EXACT or REFINE
• NLRI: prefix
• When: IMMEDIATE or DEFER
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-12

The content of the ORF value is determined by the ORF-type setting. An ORF type of NLRI-based filtering (type 1) uses the following actions:

- ADD: adds a line to a prefix-list filter on the remote peer
- DELETE: removes a line from a filter that was previously installed on a remote peer
- DELETE ALL: removes all previously installed filters on the remote peer

For each filter entry, there is a match component that specifies either PERMIT or DENY. A PERMIT asks the peer to send updates with routes that match the set of entries as specified in the outbound route filter. DENY specifies that the remote peer should not send updates for the entries matching those specified in the ORF.

For prefixes specified with a match component of PERMIT, the remote peer is asked to pass a prefix with a scope of EXACT (an exact match) or REFINE (its subnets).

Also contained within the ORF message is the when-to-refresh field. A router can set this field to IMMEDIATE (asking the remote peer to refresh as soon as it has finished processing the ORF message) or DEFER (asking the remote peer to wait until it receives a subsequent route refresh message with the same AFI/SAFI).

Practice

Q1) What ORF type is currently supported on Cisco routers?

A) 1

B) 2

C) 3

D) 128

Q2) What are three key components of an ORF type 1 message? (Choose three.)

A) action clause

B) match clause

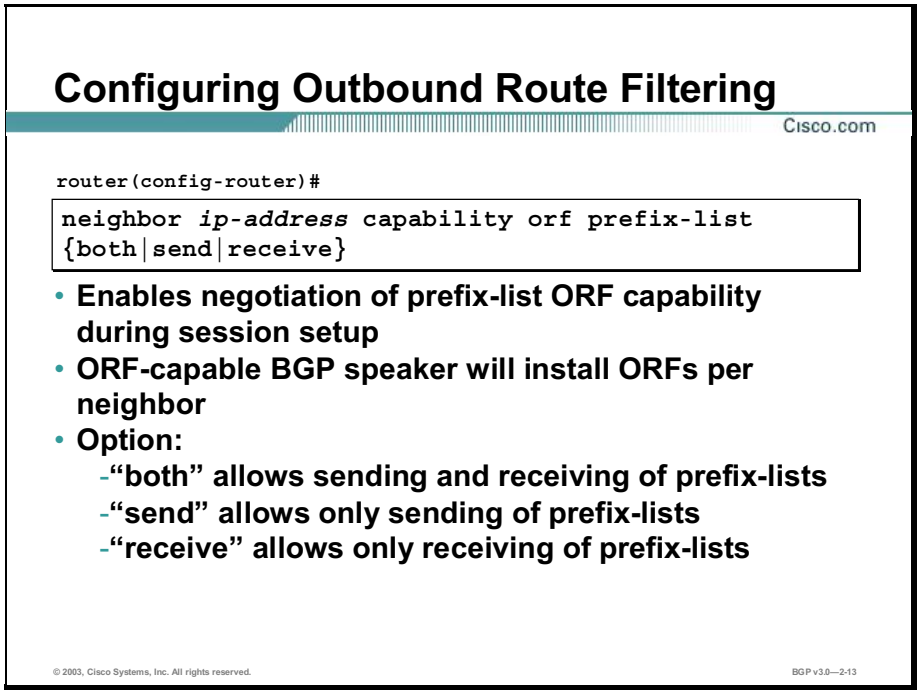
C) set clause

D) when clause

 Copyright © 2003, Cisco Systems, Inc.

Configuring Outbound Route Filtering

This topic identifies the Cisco IOS commands required to enable ORF negotiations and activate an ORF prefix-list.


Configuring Outbound Route Filtering
Cisco.com
router(config-router)#
neighbor ip-address capability orf prefix-list {both|send|receive}
• Enables negotiation of prefix-list ORF capability during session setup
• ORF-capable BGP speaker will install ORFs per neighbor
• Option:
-"both" allows sending and receiving of prefix-lists
-"send" allows only sending of prefix-lists
-"receive" allows only receiving of prefix-lists
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-13

Cisco routers support the uploading of their prefix-lists to a neighbor. You need to use the **neighbor** *ip-address* **capability orf prefix-list receive** command to advertise this capability, and you need to use **neighbor** *ip-address* **capability orf prefix-list send** to upload the inbound prefix filter to the neighbor. The uploaded filter is then used on the neighboring router after a statically configured outbound prefix-list (if it exists) is applied.

The command **neighbor** *ip-address* **capability orf prefix-list** enables the negotiation of the prefix-list ORF capability during BGP session setup. The prefix-list-based ORF (ORF type=128) is the only ORF type that Cisco IOS software supports.

neighbor orf prefix-list

To advertise ORF capabilities to a peer router, use the **neighbor orf prefix-list** command in address family or router configuration mode.

neighbor {*ip-address*} [**capability**] **orf prefix-list** [**receive** | **send** | **both**]

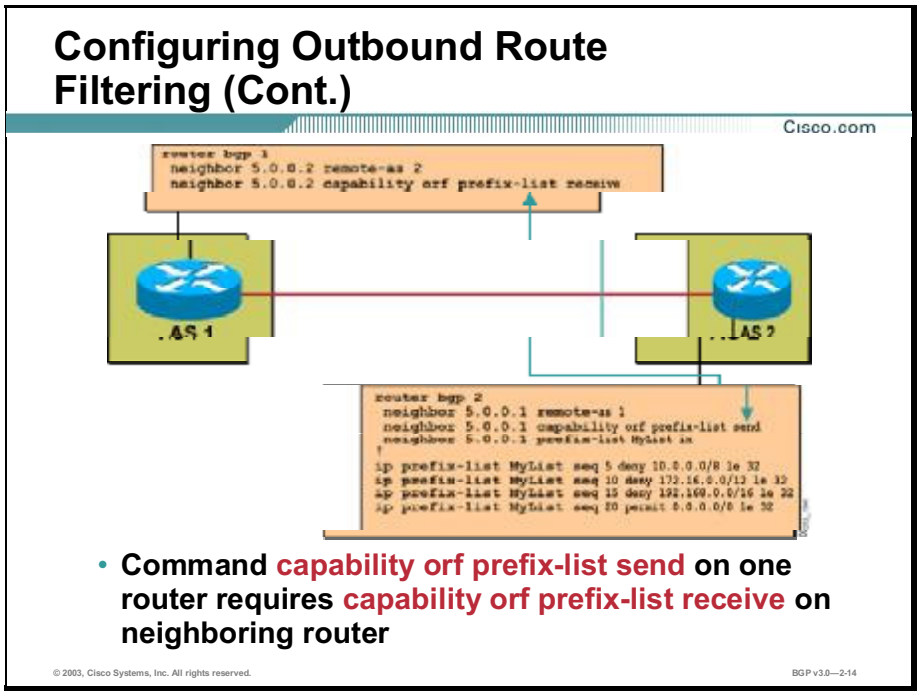
To disable ORF capabilities, use the **no** form of this command.

no neighbor {*ip-address*} [**capability**] **orf prefix-list** [**receive** | **send** | **both**]

Syntax Description

ip-address	The IP address of the neighbor router
capability	(Optional) Informs the specified neighbor that this router has ORF capabilities
receive	(Optional) Enables the ORF prefix-list capability in receive mode
send	(Optional) Enables the ORF prefix-list capability in send mode
both	(Optional) Enables the ORF prefix-list capability in both receive and send modes

 Copyright © 2003, Cisco Systems, Inc.


Configuring Outbound Route Filtering (Cont.)
Cisco.com
router bgp 1
neighbor 5.0.0.2 remote-as 2
neighbor 5.0.0.2 capability orf prefix-list receive
AS 1
AS 2
router bgp 2
neighbor 5.0.0.1 remote-as 1
neighbor 5.0.0.1 capability orf prefix-list send
neighbor 5.0.0.1 prefix-list MyList in
!
ip prefix-list MyList seq 5 deny 10.0.0.0/8 le 32
ip prefix-list MyList seq 10 deny 172.16.0.0/12 le 32
ip prefix-list MyList seq 15 deny 192.168.0.0/16 le 32
ip prefix-list MyList seq 20 permit 0.0.0.0/0 le 32
• Command capability orf prefix-list send on one router requires capability orf prefix-list receive on neighboring router
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-14

The example shows the configuration of two routers where one router has uploaded an input prefix-list to the neighbor to be used as an output filter.

The following configuration steps are necessary to enable outbound route filtering:

Step 1 Enable negotiation of outbound filtering based on prefix-lists.

Step 2 Attach an input prefix-list to a neighbor.

Step 3 Enable sending of input prefix-list to the neighbor.

Practice

Q1) How is capability negotiation for prefix-list–based ORF enabled?

A) by configuring a prefix-list filter on the router and enabling BGP soft route refresh

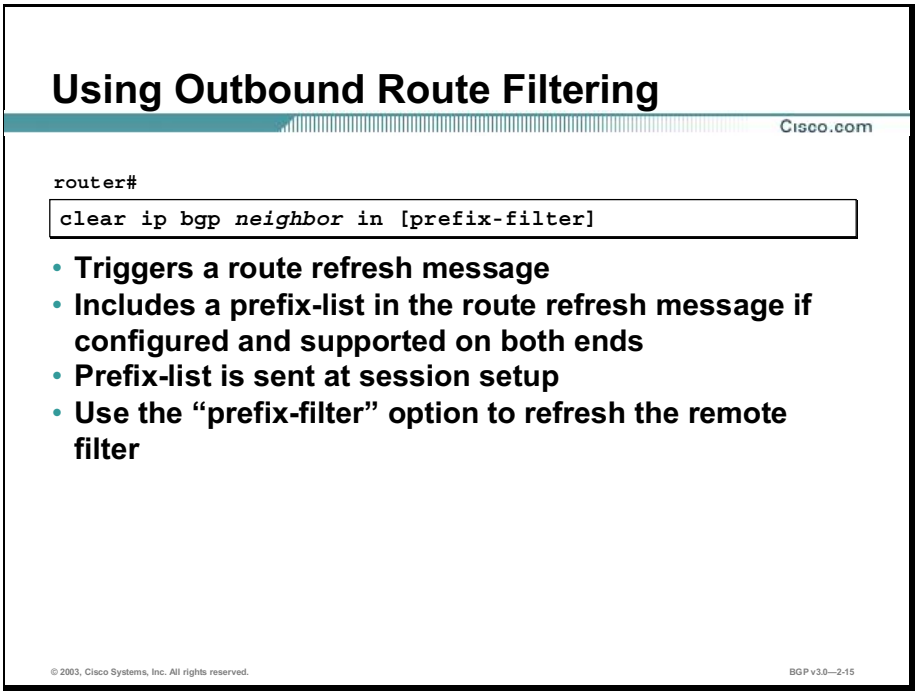
B) by issuing a **clear ip bgp neighbor** command on the router

C) by configuring both neighbors with complementing ORF capabilities using the **neighbor capability orf prefix-list** command

D) by applying a prefix-list to the BGP session with the **neighbor prefix-list** command

Using Outbound Route Filtering

This topic identifies the Cisco IOS commands used to trigger a route refresh message.


Using Outbound Route Filtering
Cisco.com
router#
clear ip bgp neighbor in [prefix-filter]
• Triggers a route refresh message
• Includes a prefix-list in the route refresh message if configured and supported on both ends
• Prefix-list is sent at session setup
• Use the "prefix-filter" option to refresh the remote filter
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-15

Use the **clear ip bgp** *neighbor* command with the **prefix-filter** keyword to push out the existing ORF prefix-list so that a new route refresh will be received from a neighbor. The neighbor will use the ORF prefix-list previously negotiated.

You need to use the **clear ip bgp** *neighbor* command only when the filter has been modified because the neighbor will store the filter for subsequent route refresh requests. The neighbor will then use the filter on all updates toward the router that originated the filter.

Note	The **in** keyword is entered when using the **clear ip bgp** *neighbor* command because inbound route refresh is desired; only the inbound prefix-list filter is pushed to the neighbor and used by the neighbor in the outbound direction.

The router will ignore the **prefix-filter** keyword if ORF capability has not been received or the send capability has not been enabled.

When the **clear ip bgp** *neighbor* command is used without the **prefix-filter** keyword, a normal route refresh is performed. The **prefix-filter** keyword should always be used when ORF inbound routing policy changes occur.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What is the recommended method of triggering an ORF route refresh?

A) Reload the remote peer router.

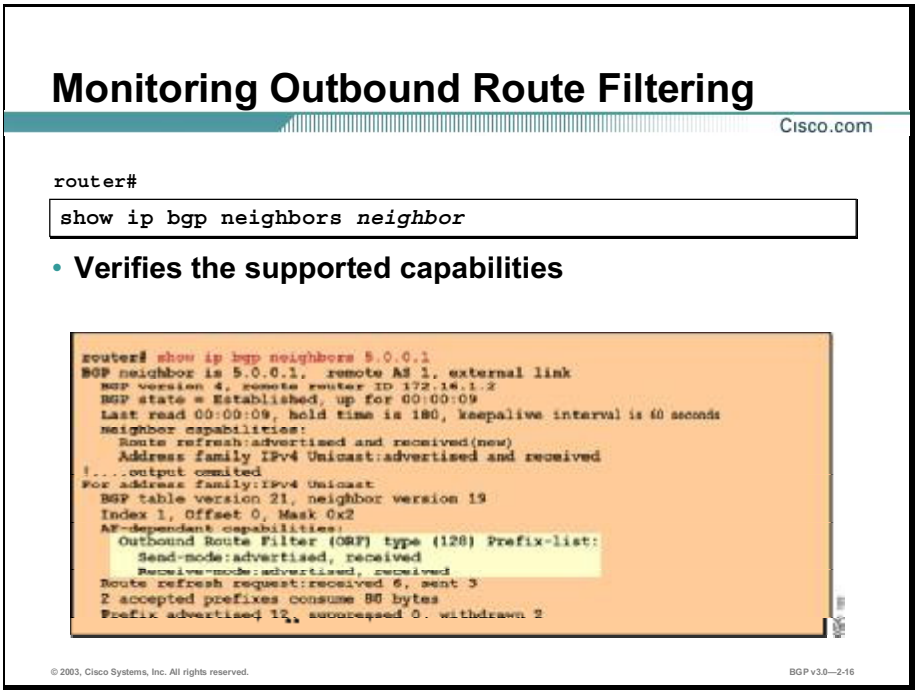
B) Remove and replace the BGP neighbor configuration.

C) Execute the **clear ip bgp *** command on the router.

D) Execute the **clear ip bgp** *neighbor* **in [prefix-filter]** command.

Monitoring Outbound Route Filtering

This topic identifies the Cisco IOS commands required to monitor the operation of an ORF that you have configured and activated.


Monitoring Outbound Route Filtering
Cisco.com
router#
show ip bgp neighbors neighbor
• Verifies the supported capabilities
router# show ip bgp neighbors 5.0.0.1
BGP neighbor is 5.0.0.1, remote AS 1, external link
BGP version 4, remote router ID 172.16.1.2
BGP state = Established, up for 00:00:09
Last read 00:00:09, hold time is 180, keepalive interval is 60 seconds
neighbor capabilities:
Route refresh:advertised and received(new)
Address family IPv4 Unicast:advertised and received
!...output omitted
For address family:IPv4 Unicast
BGP table version 21, neighbor version 19
Index 1, Offset 0, Mask 0x2
AF-dependant capabilities:
Outbound Route Filter (ORF) type (128) Prefix-list:
Send-mode:advertised, received
Receive-mode:advertised, received
Route refresh request:received 6, sent 3
2 accepted prefixes consume 80 bytes
Prefix advertised 12, suppressed 0, withdrawn 2
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-16

Use the **show ip bgp neighbors** *neighbor* command to display the supported capabilities.

If the neighbor supports a certain ORF capability, it is shown as advertised, received and a filter of the supported type can be sent by that router to its neighbor.

The example output from the **show ip bgp neighbors** command shows that neighbor 5.0.0.1 is configured with the prefix-based ORF feature in both send and receive modes. ORF capabilities negotiation has been completed and is displayed per address family. The ORF type negotiated by this router with its peer is 128 (Cisco proprietary, prefix-list-based).

Practice

Q1) What command can you use to verify that a router supports the exchange of ORF capabilities?

A) **show ip bgp summary**

B) **show cdp neighbors**

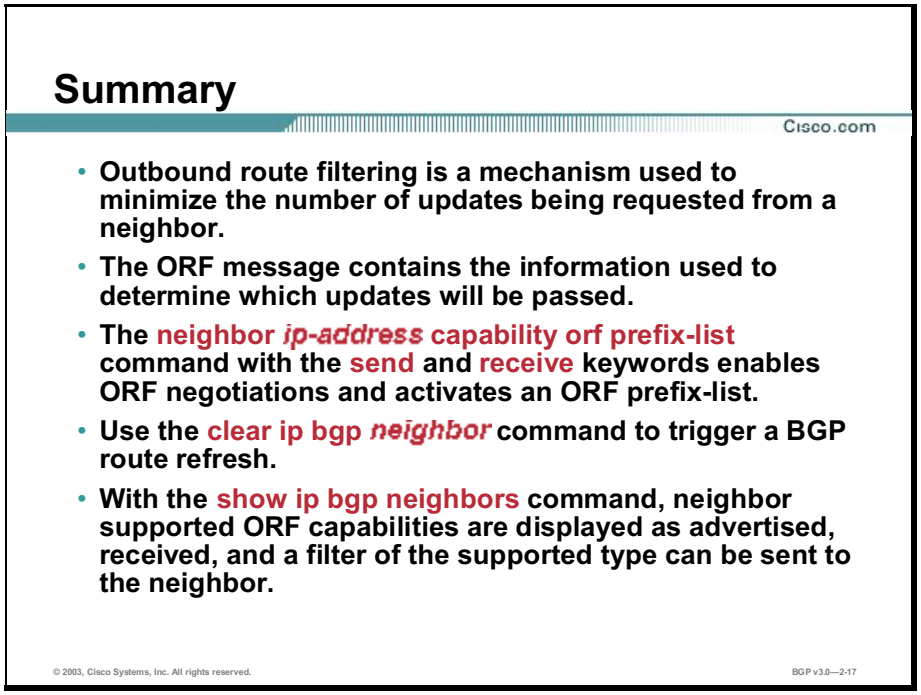
C) **show ip bgp neighbors**

D) **show ip bgp orf capabilities**

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Outbound route filtering is a mechanism used to minimize the number of updates being requested from a neighbor.
• The ORF message contains the information used to determine which updates will be passed.
• The neighbor ip-address capability orf prefix-list command with the send and receive keywords enables ORF negotiations and activates an ORF prefix-list.
• Use the clear ip bgp neighbor command to trigger a BGP route refresh.
• With the show ip bgp neighbors command, neighbor supported ORF capabilities are displayed as advertised, received, and a filter of the supported type can be sent to the neighbor.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-17

Next Steps

After completing this lesson, go to:

- Route-Maps as BGP Filters lesson

References

For additional information, refer to these resources:

- For more information on outbound route filtering, refer to “BGP Prefix-Based Outbound Route Filtering” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t11/ft11borf.htm

- For more information on ORF messages, refer to “Cooperative Route Filtering Capability for BGP-4” in the Internet Engineering Task Force (IETF) working group.

Quiz: Outbound Route Filtering

Complete the quiz to assess what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this quiz:

- The material contained within this lesson

Objectives

This quiz tests your knowledge on how to:

- Identify the function of outbound route filtering in a BGP network
- Describe the format and function of the ORF message
- Identify the Cisco IOS commands required to enable ORF negotiations and activate an ORF prefix-list
- Identify the Cisco IOS commands used to trigger a route refresh
- Identify the Cisco IOS commands required to monitor the operation of a configured ORF

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What best describes the capabilities of the proprietary ORF type supported on Cisco routers?

A) standard BGP communities filtering

B) extended BGP communities filtering

C) AS-path filtering

D) prefix-list filtering

 Copyright © 2003, Cisco Systems, Inc.

Q2) What are two key benefits to using outbound route filtering? (Choose two.)

A) conserves CPU cycles

B) improves security

C) reduces bandwidth used by unnecessary routing updates

D) increases neighbor availability

Q3) How should you configure the **neighbor capability orf prefix-list** on a router applying a prefix-list filter as an outbound route policy?

A) send

B) receive

C) both

D) prefix-filter

Q4) What are two methods of determining that a router has ORF capabilities exchange configured? (Choose two.)

A) with the **show running-config | begin bgp** command

B) using the **show ip bgp negotiate** command

C) by executing the **show ip bgp neighbors** command

D) by using the **show ip prefix-list** command

Q5) What are two prerequisites before you can configure ORF prefix-list functionality? (Choose two.)

A) A route refresh must be sent using the **clear ip bgp** command.

B) A BGP peering session between the ORF routers must be up and running.

C) ORF capabilities must be enabled on both routers.

D) You must configure a prefix-list filter on the receiving router.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Route-Maps as BGP Filters

Overview

Border Gateway Protocol (BGP) is a powerful routing protocol that supports a wide variety of administrative policy controls and route selection features. Complex filtering policies may contain requirements to match routing updates against a set of differing criteria that cannot be facilitated by a single-purpose filtering mechanism such as an AS-path filter. Route-maps provide a method to perform a variety of compound, complex filtering operations within a single tool. This lesson describes route-maps and how you can use them for BGP filtering. Included in this lesson are the commands required to use route-maps with prefix-lists, how to use them as BGP filters, and how to monitor previously configured route-maps.

Importance

Many complex filtering goals and administrative policies cannot be achieved by using only single-purpose filtering methods or compounding multiple filtering methods together. Route-maps are an important tool that can assist network operators in solving complex filtering requirements. Understanding the operation and use of route-maps is a critical component in the successful implementation of any large-scale BGP deployment.

Objectives

Upon completing this lesson, you will be able to:

- Identify the need to use route-maps to influence route selection in a BGP network
- Identify the high-level function of a route-map
- Identify the Cisco IOS® commands required to configure a route-map to match against a prefix-list
- Identify where you can apply route-maps as route filters in a BGP network

- Identify the Cisco IOS commands required to enable a route-map as a BGP route filter
- Identify the Cisco IOS commands required to monitor the operation of a configured route-map used as a BGP filter

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

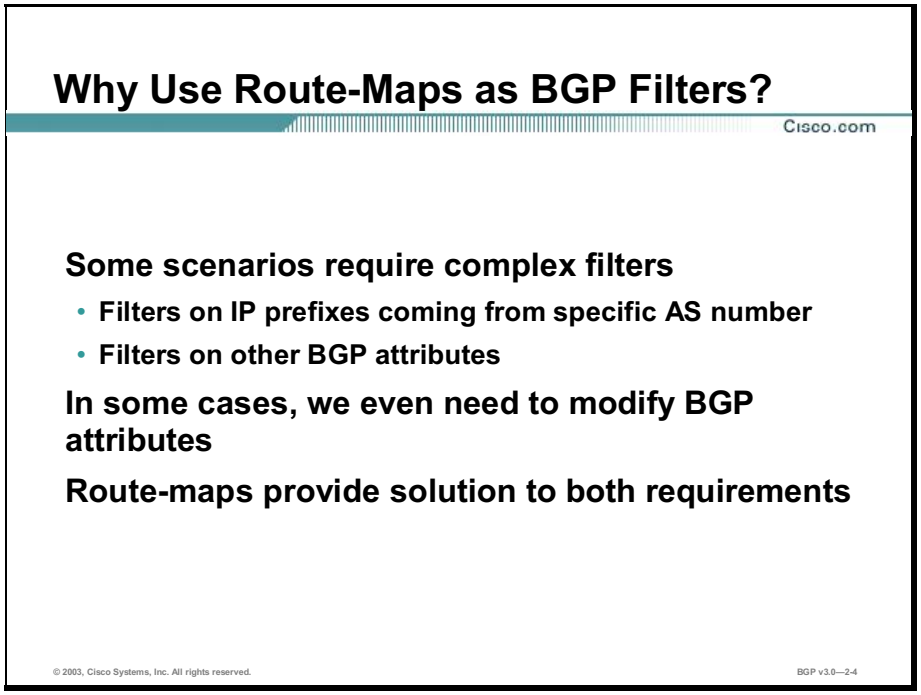
This lesson includes these topics:

- Overview
- Why Use Route-Maps as BGP Filters?
- Route-Maps Overview
- Prefix-List Use in Route-Maps
- BGP Filters
- Using Route-Maps as BGP Filters
- Monitoring Route-Maps
- Summary
- Assessment (Quiz): Route-Maps as BGP Filters

 Copyright © 2003, Cisco Systems, Inc.

Why Use Route-Maps as BGP Filters?

This topic identifies the need to use route-maps to influence route selection in a BGP network.


Why Use Route-Maps as BGP Filters?
Cisco.com
Some scenarios require complex filters
• Filters on IP prefixes coming from specific AS number
• Filters on other BGP attributes
In some cases, we even need to modify BGP attributes
Route-maps provide solution to both requirements
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-4

Network administrators cannot achieve certain complex filtering goals by using a prefix-list only or by using an AS-path filter list only. Using both of these filters simultaneously means that a route must be permitted by both to be accepted. Sometimes the goal is to permit a specific prefix if it is received with a specific AS-path and to deny it otherwise.

Combinations of tests can be implemented using route-maps. A route-map is a powerful filtering tool that can also modify routing information. Different attributes can have their values set or changed by the route-map.

Practice

Q1) What are two reasons for using route-maps on BGP neighbors? (Choose two.)

A) when you want to change the attributes of BGP updates

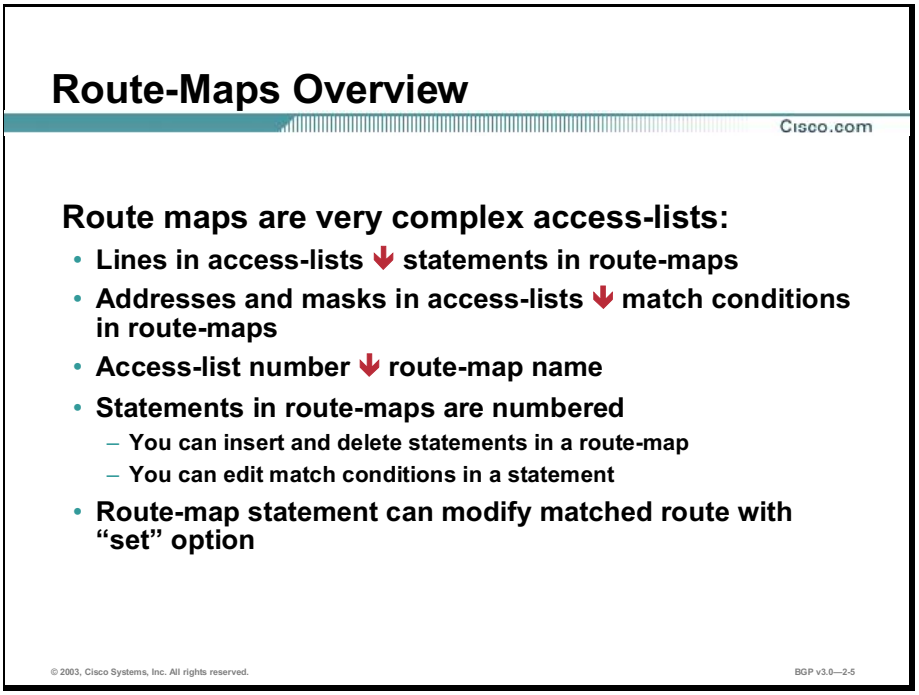
B) when you want to filter BGP updates based on a complex set of conditions

C) when you want to implement a simple routing policy for incoming and outgoing BGP updates

D) when you want to use both a prefix-list and an AS path to filter BGP updates

Route-Maps Overview

This topic identifies the high-level function of a route-map.


Route-Maps Overview
Cisco.com
Route maps are very complex access-lists:
• Lines in access-lists ↓ statements in route-maps
• Addresses and masks in access-lists ↓ match conditions in route-maps
• Access-list number ↓ route-map name
• Statements in route-maps are numbered
– You can insert and delete statements in a route-map
– You can edit match conditions in a statement
• Route-map statement can modify matched route with "set" option
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-5

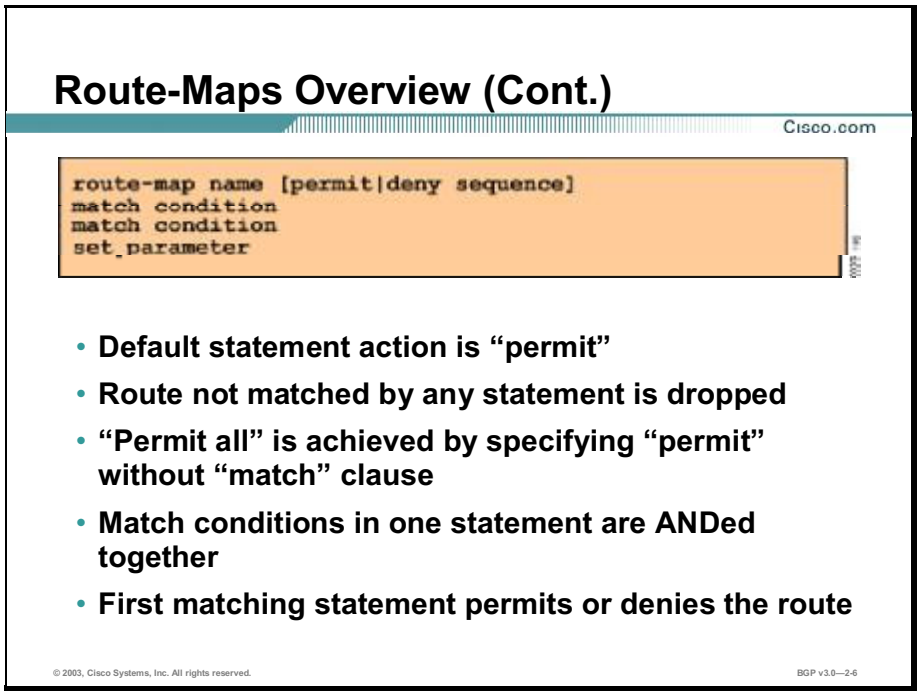
A route-map is a filter. What is denied by the route-map is dropped. Additionally, you can use the route-map to modify attributes of the permitted routes.

Route-maps have similarities to access-lists. Both have a set of tests that will be performed. Several tests can be done in sequence. The first match produces the result of either "permit" or "deny."

An access-list has a number of lines, each indicating a testing condition. The route-map is more complex than the access-list. It consists of several groups of configuration lines. Each group is called a statement. The statement has a sequence number that provides the opportunity to remove or modify an explicit statement without removing the entire route-map. It also provides the opportunity to add a new statement between two existing statements.

Each route-map statement starts with a configuration line indicating the name of the route-map, the sequence number, and whether the result should be permitted or denied if the testing matches. The statement then continues on, following configuration lines with the match clauses. Matching can be done in several ways. It could be a test on the prefix, the AS-path, or some other attribute. The statement concludes with the optional set statements, where attributes may be modified or set.

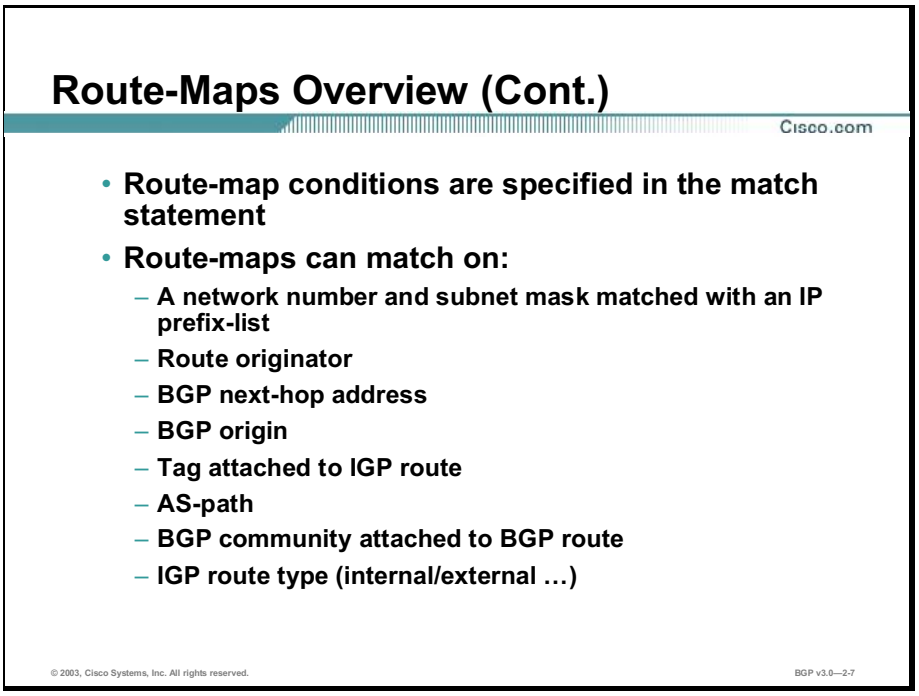
 Copyright © 2003, Cisco Systems, Inc.


Route-Maps Overview (Cont.)
Cisco.com
route-map name [permit|deny sequence]
match condition
match condition
set parameter
• Default statement action is “permit”
• Route not matched by any statement is dropped
• “Permit all” is achieved by specifying “permit” without “match” clause
• Match conditions in one statement are ANDed together
• First matching statement permits or denies the route
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-6

A route-map consists of several statements. Each statement starts with the route-map configuration line, on which the name of the route-map must be indicated. A good practice is to always indicate the **permit** or **deny** keyword followed by a sequence number.

The matching clauses for the statements are listed on the match lines following the route-map line. There may be several match lines, each referring to a different test to be performed. All tests must be passed in order for the statement to be matched. If any of the match line tests fails, the next route-map statement is tried. Statements are tried in sequence number order. If there are no more statements in the route-map, the result is, implicitly, “deny.”

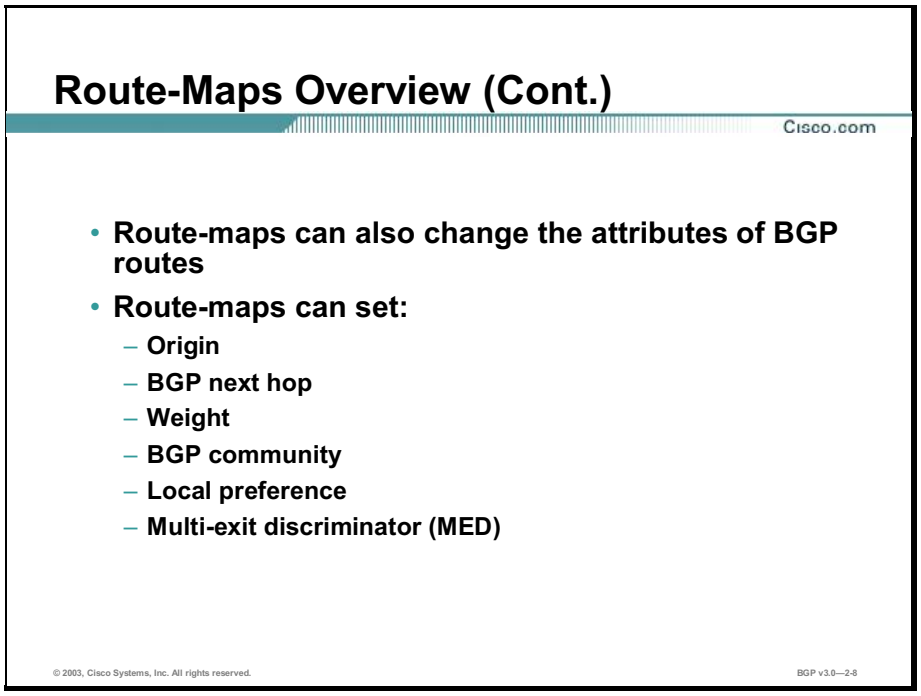
If all of the match clauses succeed, there is a match for the statement and the indicated result is used. If the result is to deny, the route is then silently ignored. If the result is to permit, the route is accepted and the set clauses are applied. The set clauses allow one or more attributes to be changed or set to specific values before the route is accepted.


Route-Maps Overview (Cont.)
Cisco.com
• Route-map conditions are specified in the match statement
• Route-maps can match on:
– A network number and subnet mask matched with an IP prefix-list
– Route originator
– BGP next-hop address
– BGP origin
– Tag attached to IGP route
– AS-path
– BGP community attached to BGP route
– IGP route type (internal/external …)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-7

Each route-map statement can have several match clauses. Each match clause is given its own configuration line. The match clause refers to the tests to be made on the candidate route. Tests of the candidate route can be based on the following criteria:

- IP network numbers and subnet masks, by referring to a prefix-list or access-list that will be applied on the route.
- Route originator, by referring to a prefix-list or access-list that will be applied on the value of the originator BGP attribute.
- Next hop, by referring to a prefix-list or access-list that will be applied on the value of the next-hop BGP attribute.
- Origin code, by testing the value of the origin BGP attribute.
- Tag value attached to an Interior Gateway Protocol (IGP) route—only used when redistribution from IGP into BGP occurs.
- AS-path, by referring to an AS-path access-list that will be applied on the value of the AS-path BGP attribute.
- Community, by referring to a community-list that will be applied on the value of the Community BGP attribute.
- IGP route type, by testing if the IGP route is internal or external—only used when redistribution from IGP to BGP occurs.

 Copyright © 2003, Cisco Systems, Inc.



Each route-map statement may have several set clauses. Each set clause is applied to the route when the route-map statement permits the route. With a route-map, the following can be set:

- Origin BGP attribute
- Next-hop BGP attribute
- Weight
- Community BGP attribute
- Local preference BGP attribute
- Multi-exit discriminator (MED) BGP attribute, by setting metric

Practice

Q1) Which BGP attributes can you match with a route-map?

A) AS-path; community; next-hop; origin; route originator

B) AS-path; community; next-hop; weight; route originator

C) AS-path; local preference; MED; next-hop; route

D) Community; MED; next-hop; origin; prefix

Q2) Which BGP attributes can you set with a route-map?

A) AS-path; community; next-hop; origin; route originator

B) AS-path; local preference; MED; next-hop; route originator

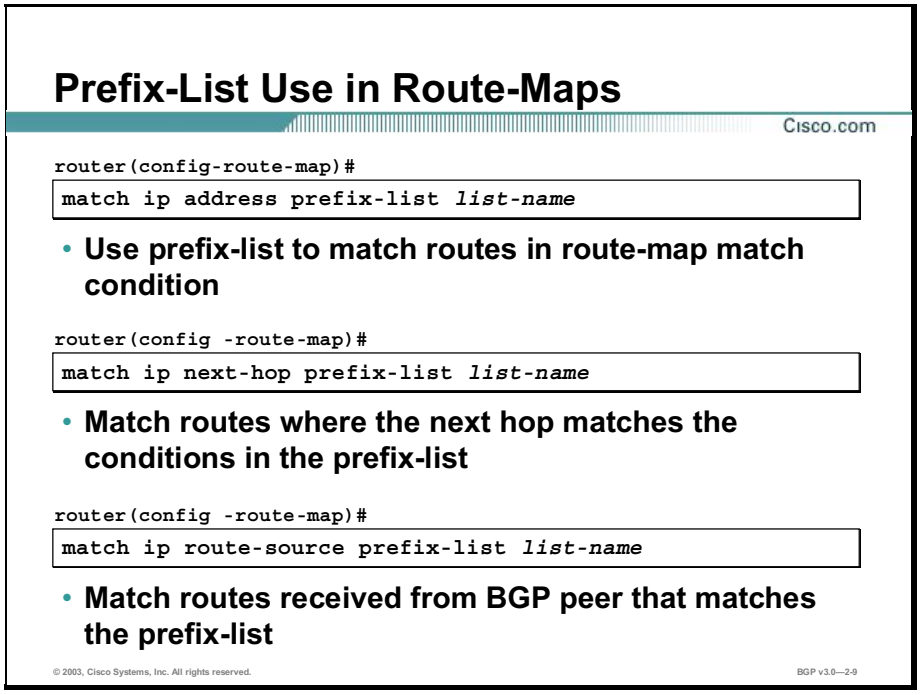
C) Community; local preference; MED; next-hop; origin

D) Community; local preference; MED; next-hop; route originator

 Copyright © 2003, Cisco Systems, Inc.

Prefix-List Use in Route-Maps

This topic identifies the Cisco IOS commands required to configure a route-map to match against a prefix-list.


Prefix-List Use in Route-Maps
Cisco.com
router(config-route-map)#
match ip address prefix-list list-name
• Use prefix-list to match routes in route-map match condition
router(config -route-map)#
match ip next-hop prefix-list list-name
• Match routes where the next hop matches the conditions in the prefix-list
router(config -route-map)#
match ip route-source prefix-list list-name
• Match routes received from BGP peer that matches the prefix-list
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-9

match ip address

To distribute any routes that have a network number that is permitted by a prefix-list, use the **match ip address** route-map configuration command. To remove the **match ip address** entry, use the **no** form of this command.

match ip next-hop

To distribute any routes that have a next-hop router address passed by one of the prefix-lists specified, use the **match ip next-hop** route-map configuration command. To remove the next-hop entry, use the **no** form of this command.

match ip route-source

To distribute routes that routers have advertised and to access servers at the address specified by the prefix-list, use the **match ip route-source** route-map configuration command. To remove the route source entry, use the **no** form of this command.

Practice

Q1) What are two ways in which route-maps are combined with prefix-lists and AS-path filters? (Choose two.)

A) You can apply route-maps, AS-path filters, and prefix-lists to incoming or outgoing routes in any combination.

B) You can use route-maps with AS-path filters but not with prefix-lists.

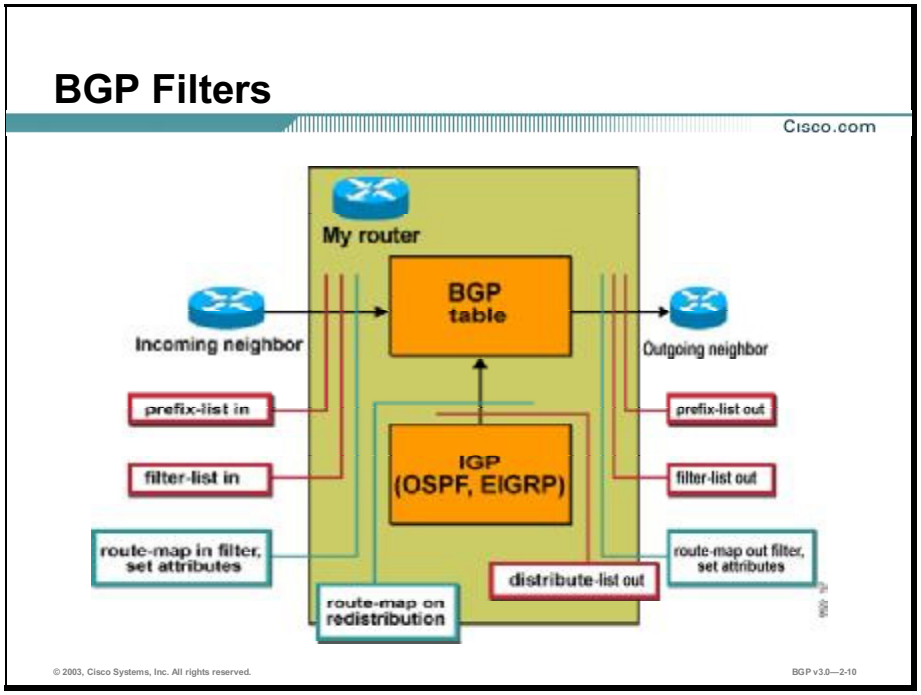
C) You can use prefix-lists and AS-path filters within the **match** statements of a route-map.

D) You cannot combine route-maps with prefix-lists or AS-path access-lists.

 Copyright © 2003, Cisco Systems, Inc.

BGP Filters

This topic identifies where you can apply route-maps as route filters in a BGP network.



You can optionally apply filter-lists, prefix-lists, and route-maps on either incoming or outgoing information, or in any combination. The incoming prefix-list, the incoming filter-list, and the incoming route-map must all permit the routes received from a neighbor before being accepted into the BGP table. Outgoing routes must pass the outgoing filter-list, the outgoing prefix-list, and the outgoing route-map before being transmitted to the neighbor.

When a router is configured to redistribute routing information from an Interior Gateway Protocol (IGP) into BGP, the routes must successfully pass any prefix-list or route-map applied to the redistribution before a route is injected into the BGP table.

Practice

Q1) What are three situations in which you can apply route-maps on a router to enable BGP policy enforcement? (Choose three.)

A) to apply policy to inbound BGP updates from a neighbor

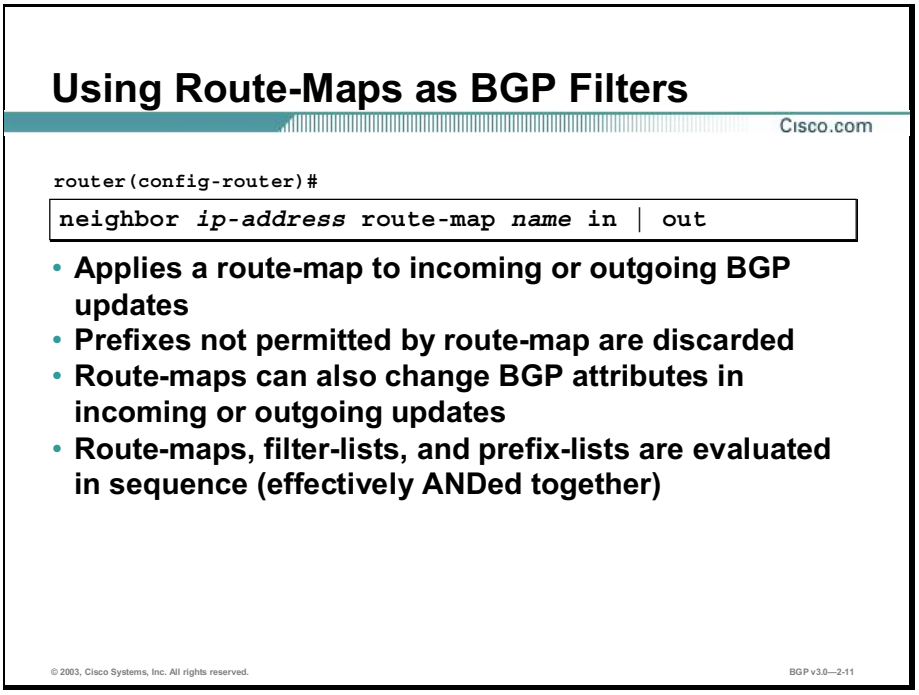
B) to apply policy to outbound BGP updates sent to a neighbor

C) to apply policy to locally originated routes

D) to filter routes distributed from an IGP into BGP

Using Route-Maps as BGP Filters

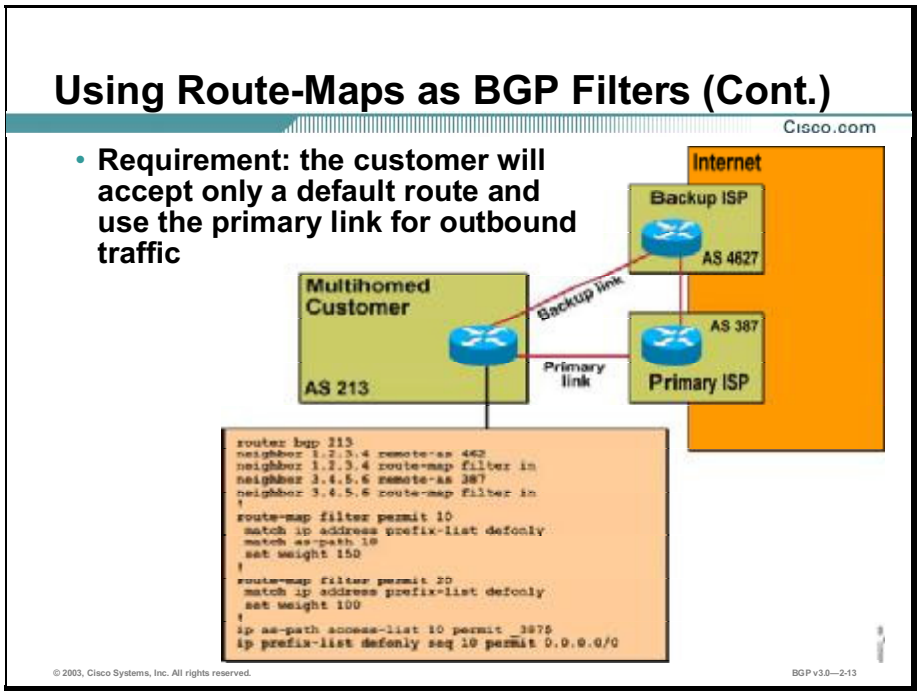
This topic identifies the Cisco IOS commands required to enable a route-map as a BGP route filter.


Using Route-Maps as BGP Filters
Cisco.com
router(config-router)#
neighbor ip-address route-map name in | out
• Applies a route-map to incoming or outgoing BGP updates
• Prefixes not permitted by route-map are discarded
• Route-maps can also change BGP attributes in incoming or outgoing updates
• Route-maps, filter-lists, and prefix-lists are evaluated in sequence (effectively ANDed together)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-11

You can apply a route-map on incoming or outgoing routing information for a neighbor. The routing information must be permitted by the route-map in order to be accepted. If there is no statement in the route-map explicitly permitting a route, then the route will be implicitly denied and dropped.

The permitted routes may have their attributes set or changed by the set clauses in the route-map. Setting attributes on routes is useful when influencing route selection. Some routes can be permitted by one of the statements in the route-map and have their attributes changed. Another statement in the route-map could permit other routes and not have their attributes altered. When route selection is performed, the attribute values indicate that one route is more preferred than the other.

 Copyright © 2003, Cisco Systems, Inc.



In this example, the customer will accept only a default route and use the primary link connected to AS 387 for outbound traffic.

Practice

Q1) How do you use a route-map with BGP?

A) Route-maps are applied globally to the BGP routing process.

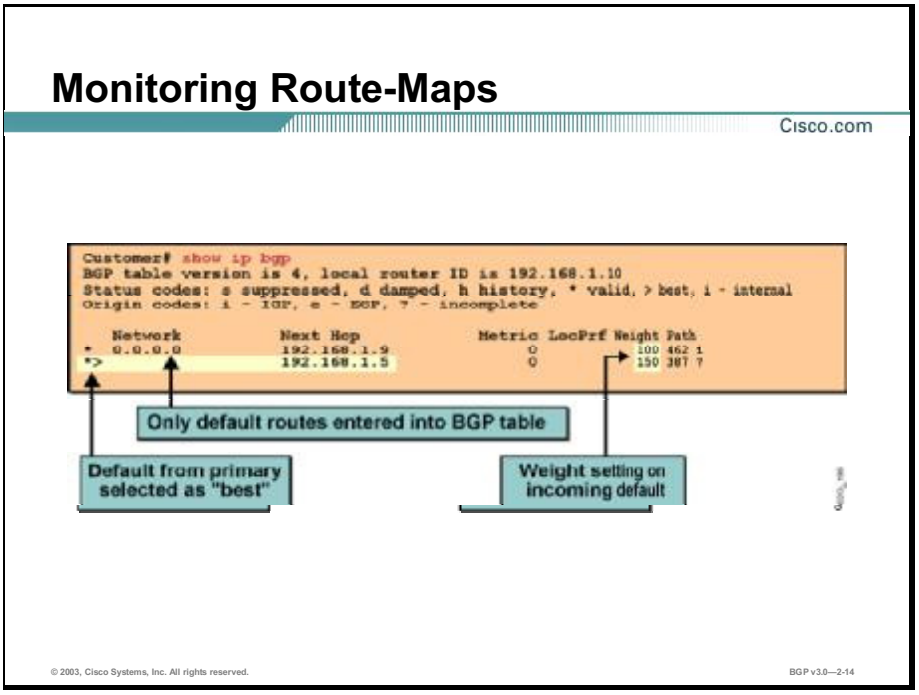
B) Route-maps are applied to BGP neighbors in the inbound or outbound direction.

C) Route-maps are applied to the interface connecting EBGP peers.

D) Route-maps are automatically appended to configured BGP neighbors.

Monitoring Route-Maps

This topic identifies the Cisco IOS commands required to monitor the operation of a configured route-map used as a BGP filter.

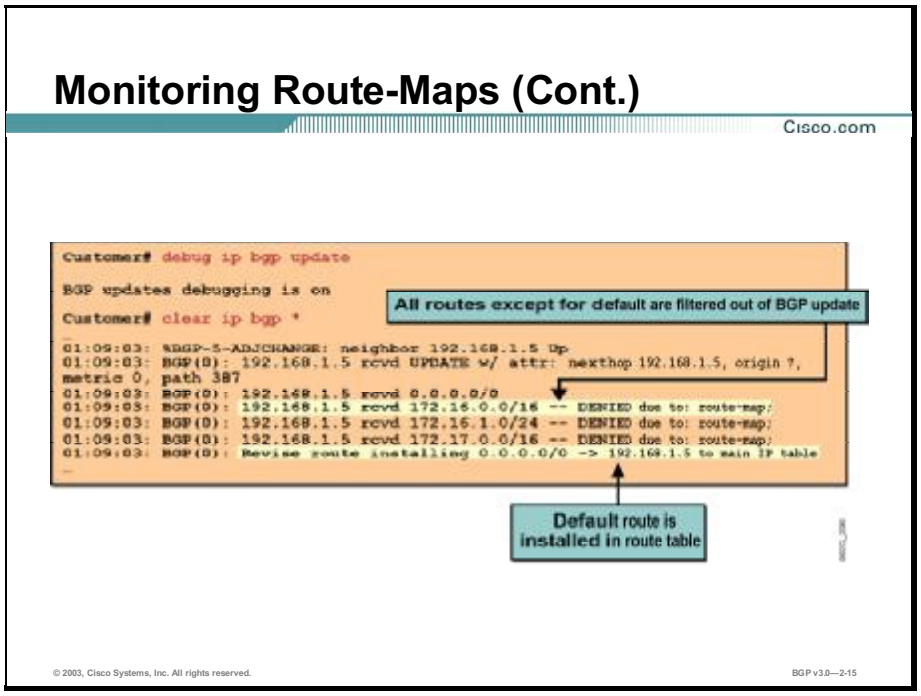

Monitoring Route-Maps
Cisco.com
Customer# show ip bgp
BGP table version is 4, local router ID is 192.168.1.10
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
* 0.0.0.0 192.168.1.9 0 100 462 1
*> 192.168.1.5 0 150 387 ?
Only default routes entered into BGP table
Default from primary selected as "best"
Weight setting on incoming default
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-14

Use the **show ip bgp** command to display the configured route-map characteristics.

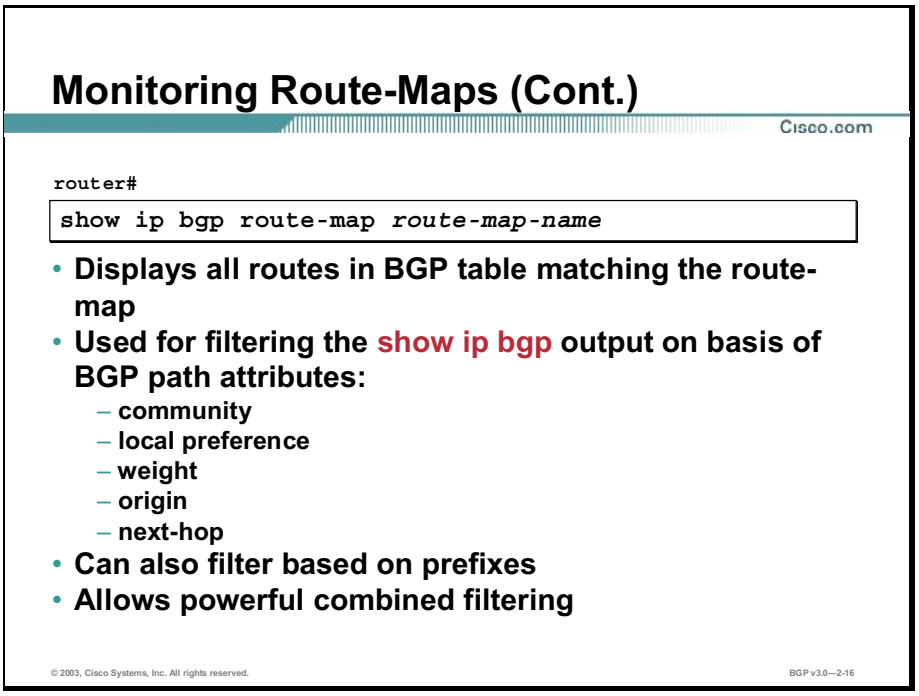
The example shows that only default routes are entered into the BGP table. The default route from the primary link has been selected by BGP as the “best” route. The BGP route selection rules have been modified based on the configuration of the BGP weight attribute in the route-map. In the configuration on the previous example, a route-map was used to set the weight of the primary link to 150 and the weight setting of the backup link to 100.

Because BGP path selection prefers the highest weight, the router uses the primary link as the outgoing path.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Route-Maps (Cont.)
Cisco.com
Customer# debug ip bgp update
BGP updates debugging is on
Customer# clear ip bgp *
01:09:03: %BGP-5-ADJCHANGE: neighbor 192.168.1.5 Up
01:09:03: BGP(0): 192.168.1.5 rcvd UPDATE w/ attr: nexthop 192.168.1.5, origin ?, metric 0, path 387
01:09:03: BGP(0): 192.168.1.5 rcvd 0.0.0.0/0
01:09:03: BGP(0): 192.168.1.5 rcvd 172.16.0.0/16 -- DENIED due to: route-map;
01:09:03: BGP(0): 192.168.1.5 rcvd 172.16.1.0/24 -- DENIED due to: route-map;
01:09:03: BGP(0): 192.168.1.5 rcvd 172.17.0.0/16 -- DENIED due to: route-map;
01:09:03: BGP(0): Revise route installing 0.0.0.0/0 -> 192.168.1.5 to main IP table
All routes except for default are filtered out of BGP update
Default route is installed in route table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-15

Here we see that all routes except for the default route are being filtered out (DENIED) of the BGP update. The default route is installed in the route table.

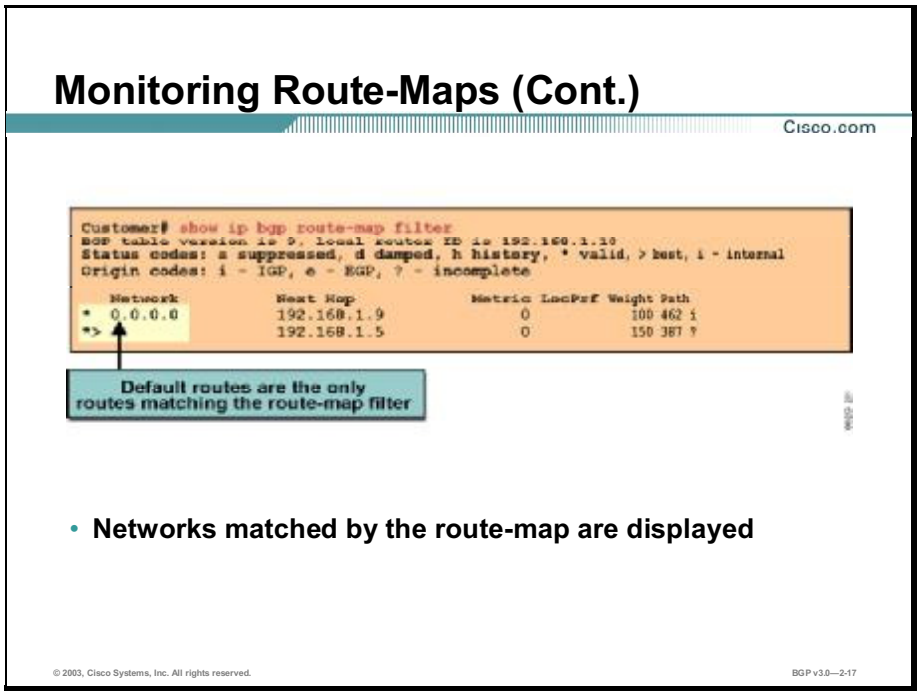

Monitoring Route-Maps (Cont.)
Cisco.com
router#
show ip bgp route-map route-map-name
• Displays all routes in BGP table matching the route-map
• Used for filtering the show ip bgp output on basis of BGP path attributes:
– community
– local preference
– weight
– origin
– next-hop
• Can also filter based on prefixes
• Allows powerful combined filtering
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-16

You can also use route-maps for selective and powerful filtering of the BGP table. The **show ip bgp route-map** command displays selected routes from a BGP routing table based on the contents of a route-map.

A route-map can match the routes on the basis of BGP path attributes (local preference, community, weight, origin, next-hop) and/or prefix-lists and access-lists (matching IP prefixes). The power of route-map filtering lays in the possibility of combining different filters; for example, filtering on community, prefix, and next-hop values.

Note	Support for route-map filtering was added in 12.2(11)T and 12.0(14)ST images of Cisco IOS software.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Route-Maps (Cont.)
Cisco.com
Customer# show ip bgp route-map filter
BGP table version is 9, local router ID is 192.168.1.10
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
* 0.0.0.0 192.168.1.9 0 100 462 i
*> 192.168.1.5 0 150 387 ?
Default routes are the only routes matching the route-map filter
• Networks matched by the route-map are displayed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-17

In this example, a customer is using a simple route-map to filter the BGP table. By using the **show ip bgp route-map** command, the customer can display the filtered BGP table. The customer router configuration from which this output is collected is shown below for reference:

```
router bgp 213
neighbor 1.2.3.4 remote-as 462
neighbor 1.2.3.4 route-map filter in
neighbor 3.4.5.6 remote-as 387
neighbor 3.4.5.6 route-map filter in
!
route-map filter permit 10
 match ip address prefix-list defonly
 match as-path 10
 set weight 150
!
route-map filter permit 20
 match ip address prefix-list defonly
 set weight 100
!
ip as-path access-list 10 permit _387$
ip prefix-list defonly seq 10 permit 0.0.0.0/0
```

The route-map “filter” matches incoming networks from two service providers. For all routes sent by the primary provider (AS 387), the local router accepts the default route only, and it is marked as the preferred route with a weight of 150. Only a default route is accepted from the backup provider, and its weight metric has been set to 100.

The customer then applies the route-map to the output of the **show ip bgp route-map** command, and only the networks conforming to the AS-path and prefix-list filters are displayed (network 0.0.0.0/0 in the example).

Practice

Q1) How does a **set** statement in an outgoing route-map affect the BGP table?

A) A copy of the modified BGP entry is stored in the BGP table but is not used.

B) The result of applying the filters to the IP routing table is applied to the BGP table.

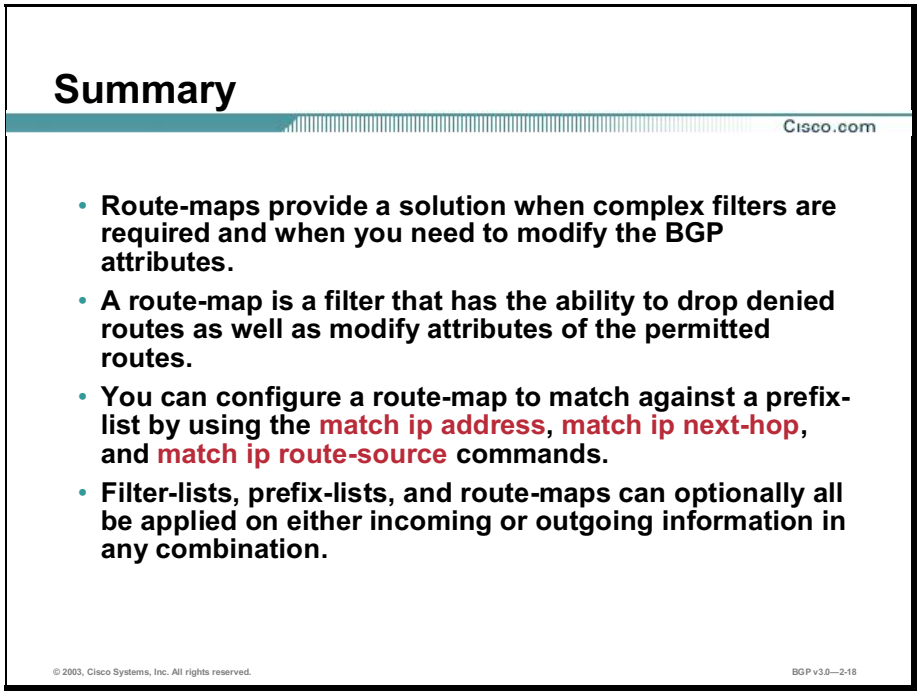
C) The changes are reflected in the BGP table.

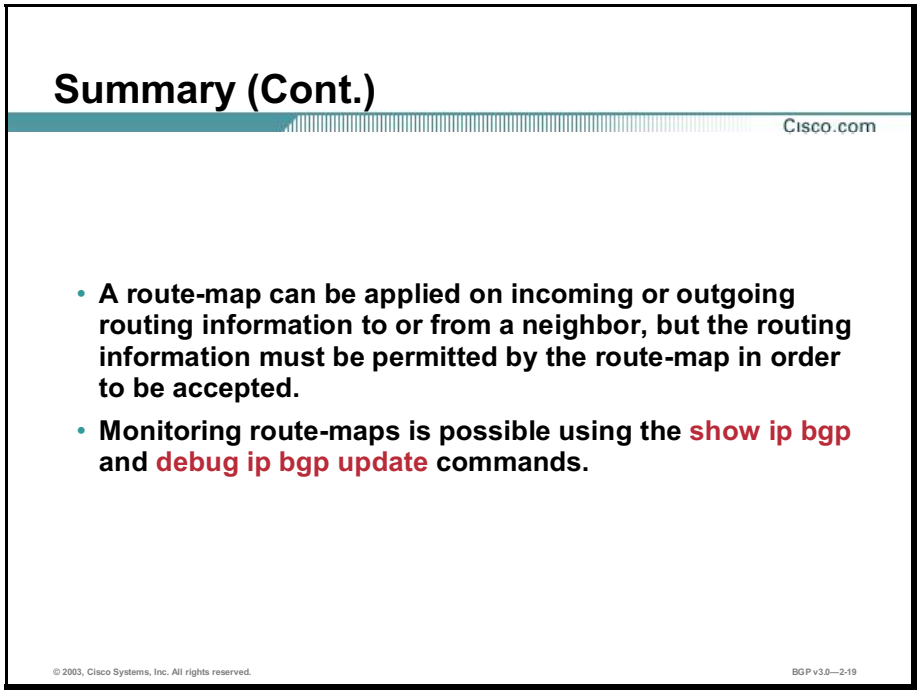
D) It does not affect the local BGP table in any way.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Route-maps provide a solution when complex filters are required and when you need to modify the BGP attributes.
• A route-map is a filter that has the ability to drop denied routes as well as modify attributes of the permitted routes.
• You can configure a route-map to match against a prefix-list by using the match ip address, match ip next-hop, and match ip route-source commands.
• Filter-lists, prefix-lists, and route-maps can optionally all be applied on either incoming or outgoing information in any combination.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-18


Summary (Cont.)
Cisco.com
• A route-map can be applied on incoming or outgoing routing information to or from a neighbor, but the routing information must be permitted by the route-map in order to be accepted.
• Monitoring route-maps is possible using the show ip bgp and debug ip bgp update commands.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-19

Next Steps

After completing this lesson, go to:

- Implementing Changes in BGP Policy lesson

References

For additional information, refer to these resources:

- For more information on route-maps, refer to “BGP Case Studies Section 1” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#routemaps

- For further information on route-maps, refer to “BGP Configuration Guide” at the following URL: http://www.cisco.com/warp/public/707/cscsupport/setup_guides/bgp.html#bgpRouteMap

 Copyright © 2003, Cisco Systems, Inc.

Quiz: Route-Maps as BGP Filters

Complete the quiz to assess what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this quiz:

- The material contained within this lesson

Objectives

This quiz tests your knowledge on how to:

- Identify the need to use route-maps to influence route selection in a BGP network
- Identify the high-level function of a route-map
- Identify the Cisco IOS commands required to configure a route-map to match against a prefix-list
- Identify where you can apply route-maps as route filters in a BGP network
- Identify the Cisco IOS commands required to enable a route-map as a BGP route filter
- Identify the Cisco IOS commands required to monitor the operation of a configured route-map used as a BGP filter

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What are three commonly used route-map match criteria in BGP environments? (Choose three.)

A) AS-path

B) prefix-list

C) community attribute

D) local preference

Copyright © 2003, Cisco Systems, Inc.

Q2) How do you implement a “permit all” when you are using route-maps?

A) By default, a route-map has an “implicit permit any” if no match is found.

B) You must configure a route-map with a “permit” parameter and no match clause.

C) You must configure a route-map with a “deny” parameter and a “deny none” clause.

D) You must configure a route-map with a “permit any” match clause.

Q3) What happens to incoming BGP updates that do not match any route-map match clauses?

A) They are entered into the BGP table.

B) They are entered into the BGP table and marked with a weight of 32768.

C) They are not accepted by the router or entered into the BGP table.

D) They are entered into the BGP table if a matching route exists in the IP routing table.

Q4) What three BGP attributes can you set using route-maps? (Choose three.)

A) MED

B) weight metric

C) next-hop

D) atomic aggregate

Q5) What are three uses of route-maps in a BGP environment? (Choose three.)

A) to filter incoming prefixes based on a prefix and the AS-path attribute

B) to modify routing information currently in the BGP table

C) to set BGP attributes such as weight and local preference on outgoing updates

D) to filter the redistribution of IGP routes into BGP

 Copyright © 2003, Cisco Systems, Inc.

Q6) What are two reasons for using route-map sequence numbers? (Choose two.)

A) to allow insertion or deletion of route-map entries

B) to order the execution sequence of route-map match clauses

C) to provide an ordered execution sequence for the route-map

D) to map between prefix-list statements and route-map match clauses

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Implementing Changes in BGP Policy

Overview

Routing policies for a Border Gateway Protocol (BGP) neighbor may include filtering mechanisms such as route-maps, distribute-lists, prefix-lists, and AS-path filter-lists. Each of these filters may impact inbound or outbound routing table updates. Whenever there is an administrative change in routing policy, the BGP session must be reset, before the new policy can take effect. To accomplish this task, there are two types of reset: hard reset and soft reset.

Clearing a BGP session using a hard reset invalidates the cache and results in a negative impact on the operation of networks, because the information in the cache becomes unavailable. A soft reset is recommended because it allows routing tables to be reconfigured and activated without clearing the BGP session.

This lesson discusses routing updates in a BGP environment and the traditional methods of forcing BGP route updates after changing a filter policy. The function and benefits of soft reconfiguration and route refresh are also discussed. The lesson also presents the commands that are required to perform a soft reconfiguration and route refresh and explains how to monitor and troubleshoot these features.

Importance

Because of the huge volumes of routing information that BGP is capable of handling, traditional routing update methods are not feasible. When an administrator changes a filter policy, the change must be calculated by the router and updated throughout the network. Three methods available to update policy changes are hard reset and soft reset (including soft reconfiguration and route refresh). A hard reset is disruptive to the BGP session, and therefore network administrators should always use the soft reset methods in production BGP networks.

Objectives

Upon completing this lesson, you will be able to:

- Identify the limitations of the traditional methods of forcing BGP route updates after changing a filter policy
- Describe the function and impact of the soft reconfiguration feature
- Identify the Cisco IOS® commands required to configure and perform a soft reconfiguration
- Identify the Cisco IOS tools available to monitor the operations of a soft reconfiguration
- Describe the function and benefits of the route refresh function
- Identify the Cisco IOS commands used to trigger a route refresh
- Identify the Cisco IOS commands required to monitor route refresh operation

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

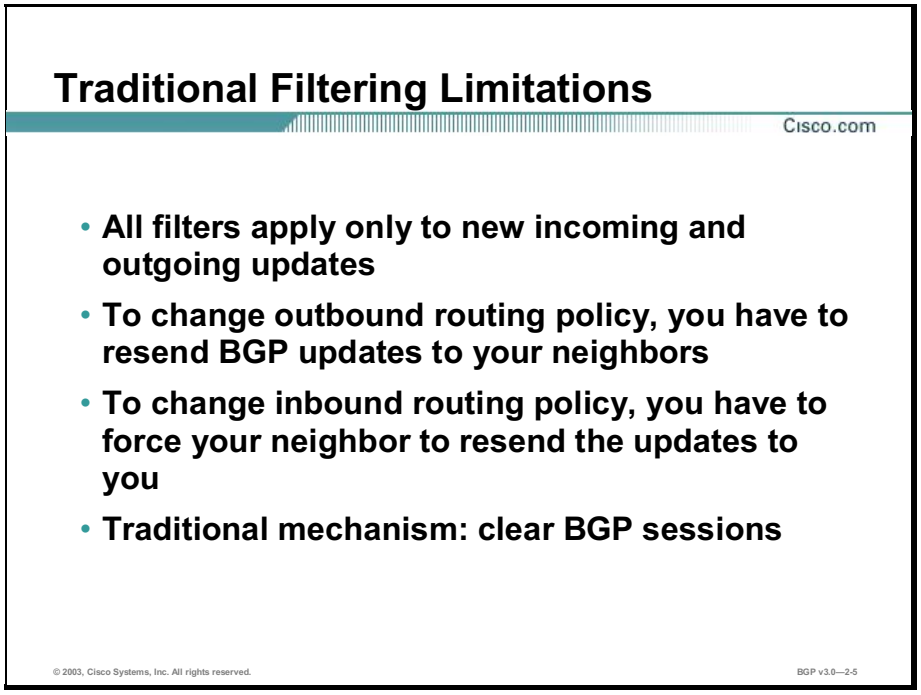
This lesson includes these topics:

- Overview
- Traditional Filtering Limitations
- BGP Soft Reconfiguration
- Cisco IOS Commands for Soft Reconfiguration
- Monitoring Soft Reconfiguration
- Route Refresh
- Using Route Refresh
- Monitoring Route Refresh
- Summary
- Assessment (Lab): Implementing Changes in BGP Policy

 Copyright © 2003, Cisco Systems, Inc.

Traditional Filtering Limitations

This topic identifies the limitations of traditional methods when forcing BGP route updates after changing filter policies.

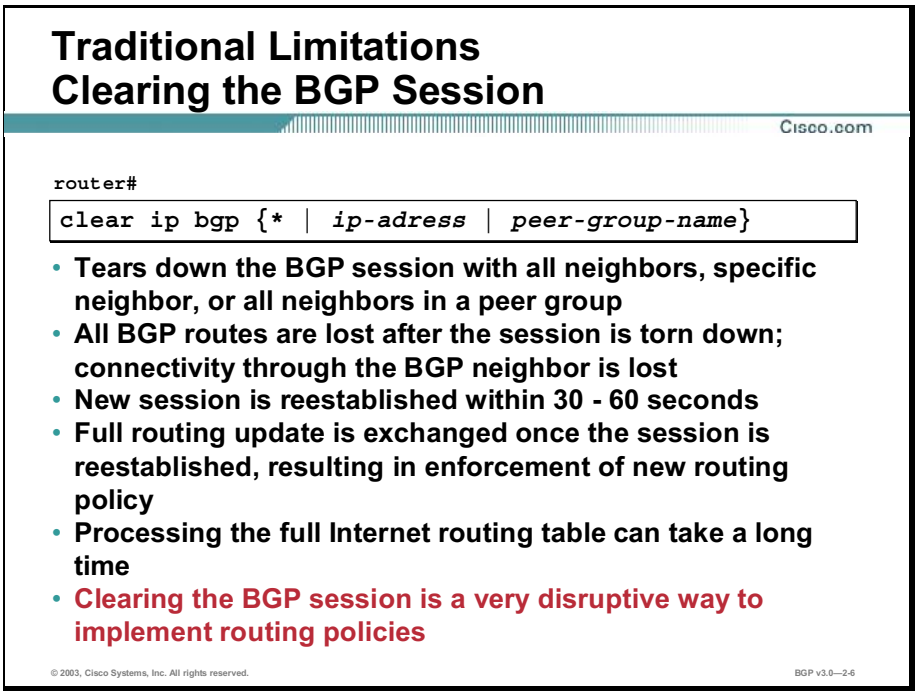

Traditional Filtering Limitations
Cisco.com
• All filters apply only to new incoming and outgoing updates
• To change outbound routing policy, you have to resend BGP updates to your neighbors
• To change inbound routing policy, you have to force your neighbor to resend the updates to you
• Traditional mechanism: clear BGP sessions
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-5

BGP can potentially handle huge volumes of routing information. But when network administrators change configuration lines in filters or route-maps, the router cannot go through the huge table of BGP information and calculate which entry is no longer valid in the local table. Nor can the router determine which route or routes, already advertised, should be withdrawn from a neighbor. There is an obvious risk that the first configuration change will be immediately followed by a second, which would cause the whole process to start all over again.

To avoid such a problem, Cisco IOS software applies changes only on those updates received or transmitted after the configuration change has been performed. This approach means that the new routing policy, enforced by the new filters, is applied only on routes received or sent after the change. If network administrators would like to apply the policy change on all routes, they have to force the router to let all routes pass through the new filter.

If the filter is applied to outgoing information, the router has to resend the entire BGP table through the new filter. If the filter is applied to incoming information, the router needs its neighbor to resend its entire BGP table so that it passes through the new filters.

Traditionally, in order to accomplish these goals, network administrators have torn down the affected BGP sessions after completing a configuration change. After the sessions are down, all information received on those sessions is invalidated and removed from the BGP table. Also, the remote neighbor will detect a session down state, and it likewise will invalidate the routes received on the session. After a period of 30 to 60 seconds, the sessions are re-established automatically and the entire BGP table is exchanged again, but through the new filters. This process, however, disrupts packet forwarding.


Traditional Limitations
Clearing the BGP Session
Cisco.com
router#
clear ip bgp {* | ip-adress | peer-group-name}
• Tears down the BGP session with all neighbors, specific neighbor, or all neighbors in a peer group
• All BGP routes are lost after the session is torn down; connectivity through the BGP neighbor is lost
• New session is reestablished within 30 - 60 seconds
• Full routing update is exchanged once the session is reestablished, resulting in enforcement of new routing policy
• Processing the full Internet routing table can take a long time
• Clearing the BGP session is a very disruptive way to implement routing policies
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-6

The EXEC command **clear ip bgp** tears down one or several BGP sessions. The BGP sessions are terminated, and the TCP connections closed. The neighbors go into the Idle state and stay there for approximately 30 seconds. Next, the neighbor session goes into the Active state, and the sessions are re-established.

You can implement the **clear ip bgp** command with the * argument, which means all sessions, or make a reference to a specific session or group of sessions to tear down.

When the session is down, all routes received over the session by both routers are invalidated. When the session is once again in the "Established" state, all BGP routes have to be resent by both peers and pass through the new filters, which enforces the new policy.

Exchanging the complete Internet routing table takes time, bandwidth, and CPU resources. IP packet forwarding to and from the neighbor is down for several minutes. Also, revoking and reannouncing the routes will be registered by the rest of the Internet as a flap for each route.

Practice

Q1) Why is clearing a BGP session a disruptive change in routing policy?

A) Clearing a BGP session takes a long time and can disrupt packet forwarding.

B) You cannot recover information sent while the BGP session is being cleared.

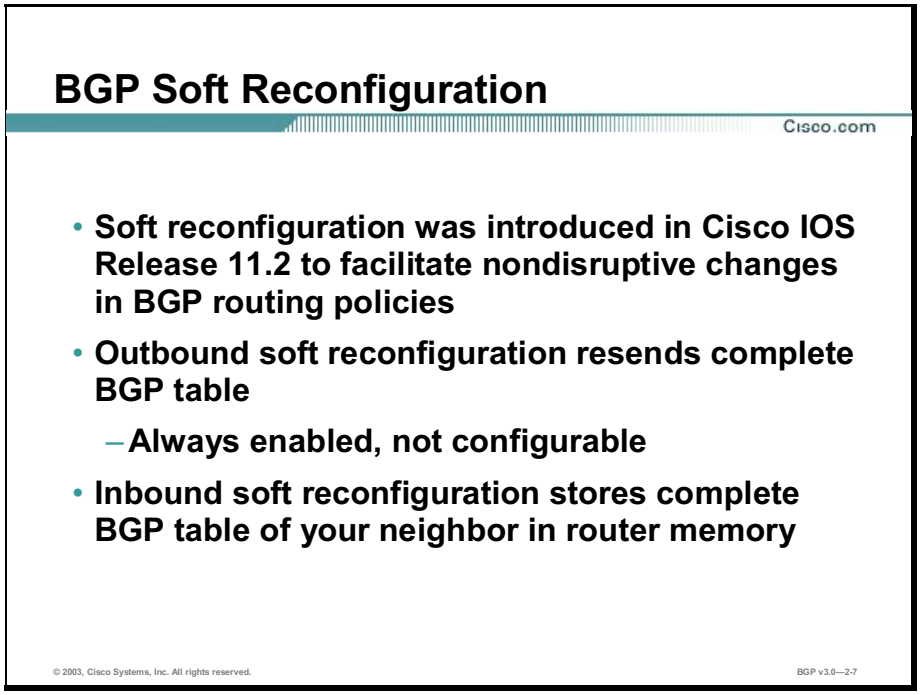
C) You cannot automatically re-establish sessions that are torn down during the clearing operation.

D) You cannot selectively tear down BGP sessions; you must clear sessions with all neighbors.

 Copyright © 2003, Cisco Systems, Inc.

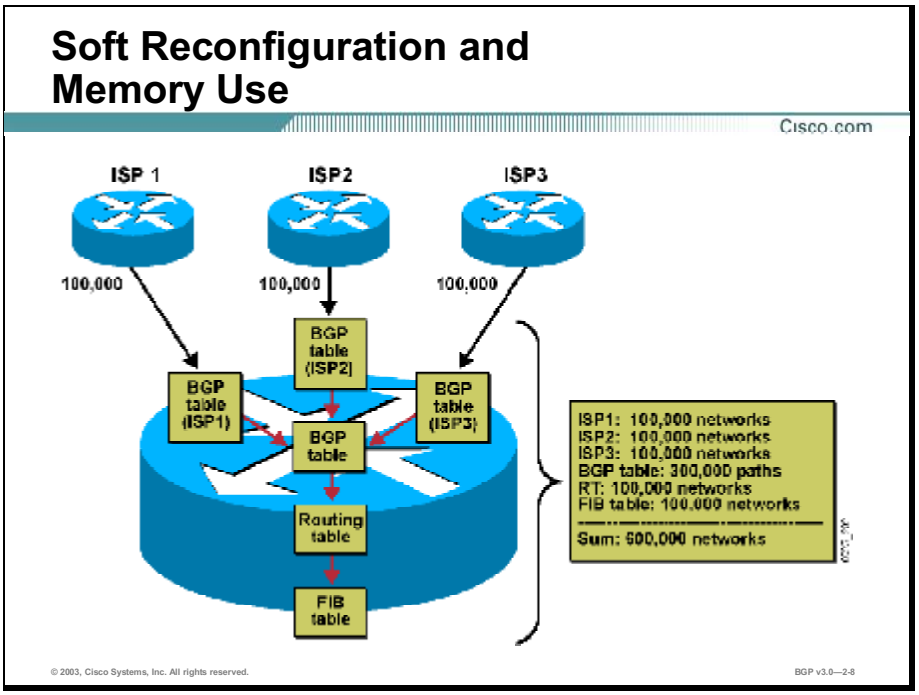
BGP Soft Reconfiguration

This topic describes the function of the soft reconfiguration feature.


BGP Soft Reconfiguration
Cisco.com
• Soft reconfiguration was introduced in Cisco IOS Release 11.2 to facilitate nondisruptive changes in BGP routing policies
• Outbound soft reconfiguration resends complete BGP table
– Always enabled, not configurable
• Inbound soft reconfiguration stores complete BGP table of your neighbor in router memory
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-7

With Cisco IOS Release 11.2 came the introduction of the soft reconfiguration feature. Soft reconfiguration provides the ability to run all routes through the filters without tearing down the sessions. Outbound soft reconfiguration was easy to implement because it is a simple resending of all routes in the local BGP table. Inbound soft reconfiguration is more complicated because a copy of all the routes received from a neighbor is required. The copy of the routes received from a neighbor is saved independent of the BGP table, before any filters are applied. Whenever the incoming filters are changed, a replay of everything received from the neighbor will take place without involving the neighbor. The major drawback of this approach is the amount of memory required to hold the copy.

Example



This example shows the impact of soft reconfiguration on an Internet service provider (ISP) router with three upstream neighbors sending full Internet routing information.

Each neighbor is sending 100,000 prefixes. The router stores each set in a dedicated per-neighbor BGP table. All 300,000 paths will then appear in the main BGP table if there is no filtering. The router will then choose the best path for each prefix and put it into the routing table. If Cisco Express Forwarding (CEF) switching is enabled, the router will store another copy in the Forwarding Information Base (FIB) table.

This solution obviously does not scale in terms of the number of neighbors and prefixes.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What is the impact of inbound soft reconfiguration?

A) It clears the session after you reconfigure the new routing policy.

B) It creates a copy of all routes received from a neighbor after the filters are applied.

C) It requires extra memory to hold a copy of all routes received from the neighbor.

D) It resets the table version number of the neighbor to 0.

Q2) What are two impacts of outbound soft reconfiguration? (Choose two.)

A) No extra CPU resources will be consumed.

B) No extra memory will be consumed.

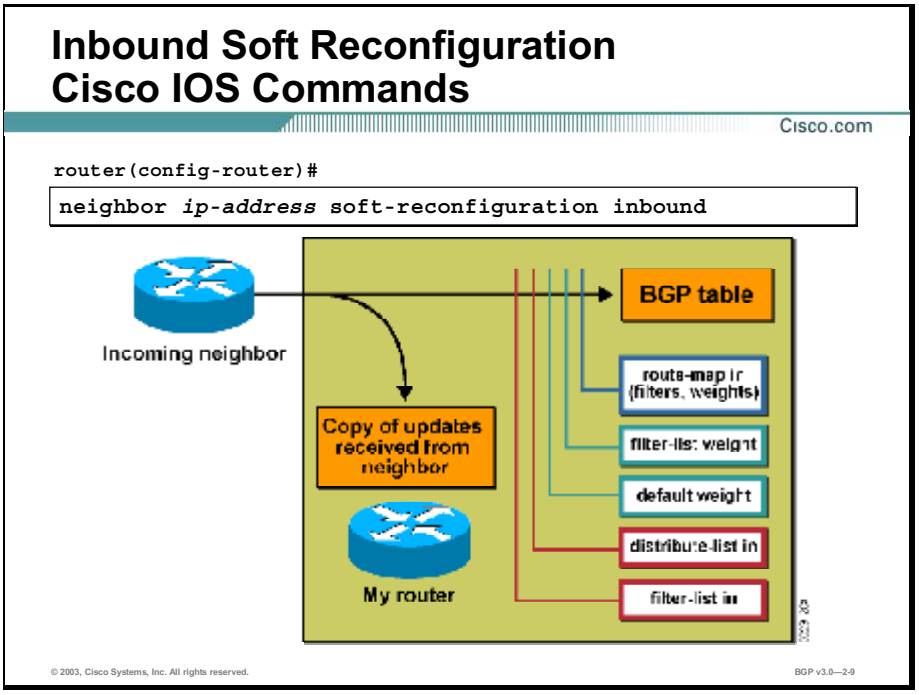
C) Extra memory is required, because a copy of the BGP table is maintained.

D) Extra CPU resources are consumed only while the **clear ip bgp** command is being issued.

Copyright © 2003, Cisco Systems, Inc.

Cisco IOS Commands for Soft Reconfiguration

This topic identifies the Cisco IOS commands required to configure and perform a soft reconfiguration.



When you configure **soft-reconfiguration inbound** for a neighbor, the router stores all routes received from that neighbor as an extra copy in memory. This copy is taken before any filtering is applied by the router to routes it receives.

This process is not enabled by default because it may consume large volumes of memory.

neighbor soft-reconfiguration

To configure Cisco IOS software to start storing updates, use the **neighbor soft-reconfiguration** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **soft-reconfiguration** [**inbound**]

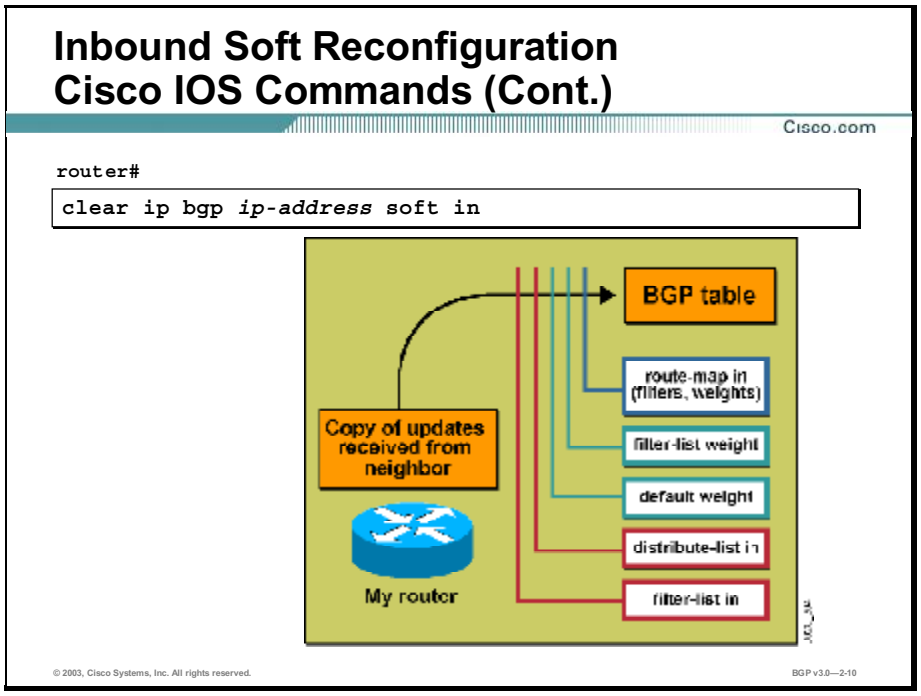
To not store received updates, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **soft-reconfiguration** [**inbound**]

Syntax Description

ip-address	IP address of the BGP-speaking neighbor
peer-group-name	Name of a BGP peer group
inbound	(Optional) Keyword indicating that the update to be stored is an incoming update

 Copyright © 2003, Cisco Systems, Inc.



When the network administrator has completed the changes to filters and route-maps that are applied on incoming information (changes that will implement a new routing policy), the **clear ip bgp** *ip-address* **soft in** is executed on the router in privileged Exec mode. Once the command has been entered, the router will not tear the session down. Instead, the router resends the saved copy of the received routing information through the new filters, and the result is stored in the local BGP table.

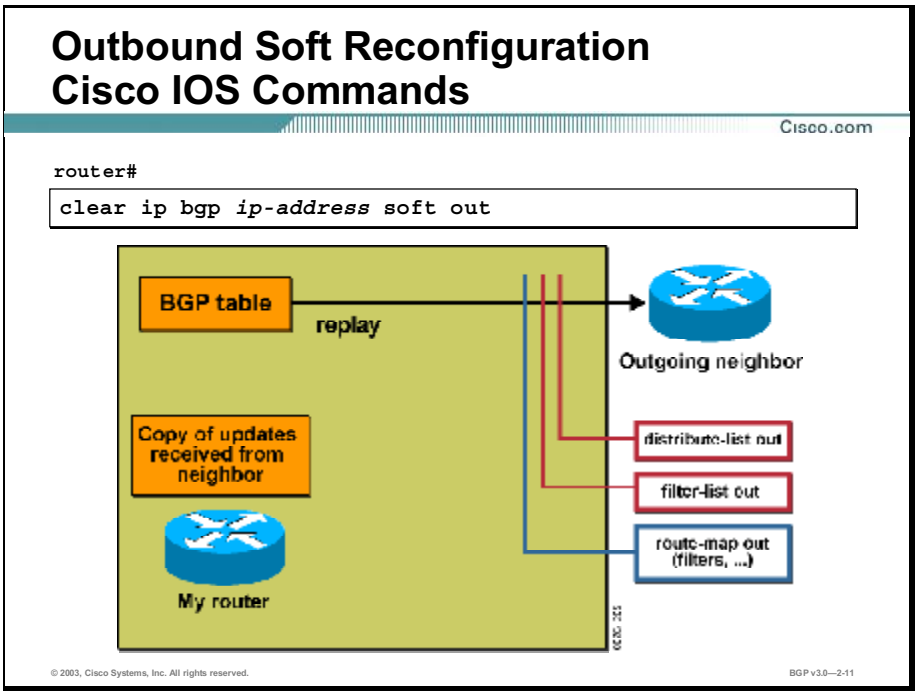
clear ip bgp

To reset a BGP connection using BGP soft reconfiguration, use the **clear ip bgp** EXEC command at the system prompt.

clear ip bgp {* | *address* | *peer-group-name*} [**soft** [**in** | **out**]]

Syntax Description

*	Resets all current BGP sessions.
address	Resets only the identified BGP neighbor.
peer-group-name	Resets the specified BGP peer group.
soft	(Optional) Soft reset. Does not reset the session.
in | **out**	(Optional) Triggers inbound or outbound soft reconfiguration. If the **in** or **out** option is not specified, both inbound and outbound soft reset are triggered.



When the network administrator has completed the changes to filters and route-maps that are applied on the outgoing information (changes that will implement a new routing policy), the **clear ip bgp** *ip-address* **soft out** is executed on the router in privileged Exec mode. Once the command has been entered, the router will not tear the session down. Instead, the table version number of the neighbor is reset to 0. When the next update interval for the neighbor arrives, the local router will go through the entire BGP table and find that all the routes need to be sent to the neighbor because they all have a table version number higher than 0.

This process causes all the BGP routes to be resent through the new filters.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What two steps must you complete to use inbound soft configuration functionality? (Choose two.)

A) Clear the BGP session inbound in the local router.

B) Clear the BGP session outbound on the remote router.

C) Configure the local neighbor with **soft-reconfiguration in**.

D) Configure the remote neighbor with **soft-reconfiguration out**.

Q2) If the **in** or **out** option is not specified when a network administrator is using soft reconfiguration, what does the router do?

A) The router will initiate a soft reconfiguration inbound.

B) The router will initiate a soft reconfiguration outbound.

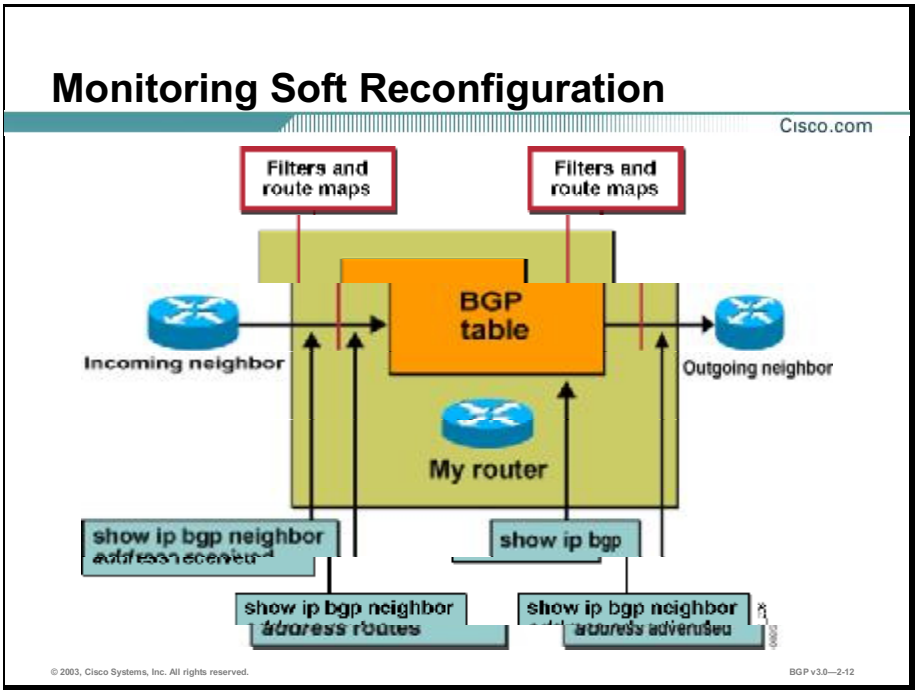
C) The router will initiate a soft reconfiguration both inbound and outbound.

D) The router will not initiate a soft reconfiguration until either the inbound or outbound option is specified.

Copyright © 2003, Cisco Systems, Inc.

Monitoring Soft Reconfiguration

This topic identifies the Cisco IOS tools available to monitor the operations of a soft reconfiguration



The **show ip bgp** command is used to display the local BGP table. You can check the entries that have been propagated to a specific neighbor with the **show ip bgp neighbor** *ip-address* **advertised** command. It displays the subset of the local BGP table that has passed the split-horizon check and all outgoing filters for the neighbor.

You can check incoming information received from a neighbor with the **show ip bgp neighbor** *ip-address* **routes** command. It displays which of the routes in the local BGP table were received (and accepted) from the indicated neighbor. Only routes passed by the incoming filter for the neighbor are displayed.

If the **soft-reconfiguration inbound** feature is enabled for a neighbor, the information saved in the extra copy outside the filters is displayed using the **show ip bgp neighbor** *ip-address* **received** command.

These commands are useful when you are troubleshooting the routing policy. You can compare routes outside the incoming filters with what was actually accepted into the BGP table from a neighbor. In addition, routes transmitted and advertised to a neighbor can be compared to what is inside the outgoing filters in the local BGP table.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) How do you check the stored BGP table when you are using inbound soft reconfiguration?

A) With the **show ip bgp neighbor** *ip-address* **advertised** command.

B) With the **show ip bgp summary route advertised** command.

C) With the **show ip bgp neighbor** *ip-address* **received** command.

D) There is no means to display the stored copy of the BGP table before soft reconfiguration.

Route Refresh

This topic describes the function and benefits of the route refresh function.

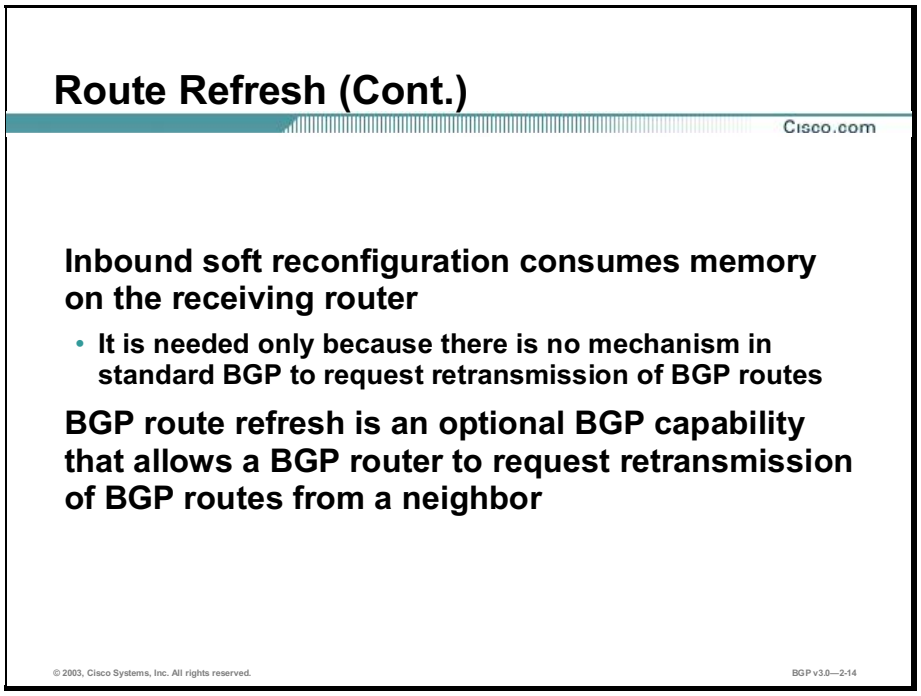
Route refresh is one of the new capabilities of BGP. Routers use the route refresh feature to request a neighbor to resend all the routing information when needed.

There are several ways of refreshing the routing information from a certain neighbor:

- Clearing the neighbor relationship
- Soft-clearing the neighbor relationship (if soft reconfiguration is enabled for this specific neighbor)
- Using route refresh (if the neighbor supports this capability)

Note	To use soft reset without preconfiguration, both BGP peers must support the soft route refresh capability, which is advertised in the OPEN message sent when the peers establish a TCP session. Routers running Cisco IOS software releases prior to release 12.1 do not support the route refresh capability and must clear the BGP session using the **neighbor soft-reconfiguration** command.

 Copyright © 2003, Cisco Systems, Inc.


Route Refresh (Cont.)
Cisco.com
Inbound soft reconfiguration consumes memory on the receiving router
• It is needed only because there is no mechanism in standard BGP to request retransmission of BGP routes
BGP route refresh is an optional BGP capability that allows a BGP router to request retransmission of BGP routes from a neighbor
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-14

The **soft-reconfiguration inbound** feature consumes large volumes of memory in the Internet environment. The number of routes received from a peer router on the Internet is so large that it is not feasible to store an extra copy.

The only reason for making the extra copy is to be able to replay the data through the new routing policy without tearing down the session and re-establishing it.

What is needed is a mechanism to ask the neighbor router to do a **clear soft outbound**. If this were possible, the extra copy would not be needed. The neighboring router, of course, has its own copy in its BGP table, which it could resend to the local router whenever signaled to.

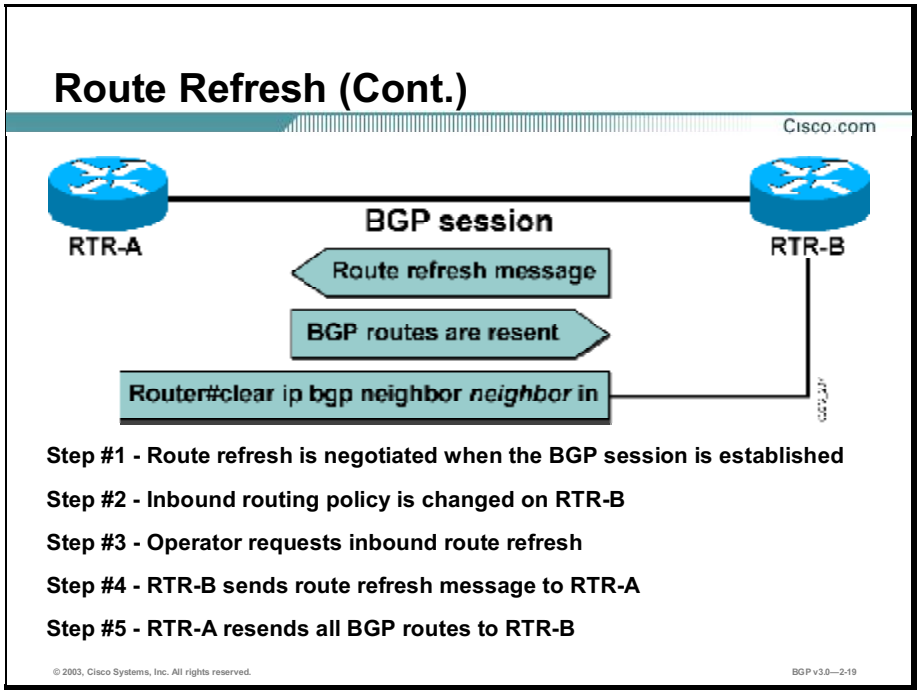
There is no such mechanism in standard BGP, but there is an optional BGP capability that allows one router to request a refresh from its neighbor: route refresh.

The following table compares the different methods of BGP session reset, stating the advantages and disadvantages of each.

Type of Reset	Advantages	Disadvantages
hard reset	No memory overhead.	The prefixes in the BGP, IP, and FIB tables provided by the neighbor are lost. Not recommended.
outbound soft reset	No configuration, no storing of routing table updates.	Does not reset inbound routing table updates.
dynamic inbound soft reset	Does not clear the BGP session or cache. Does not require storing of routing table updates, and has no memory overhead.	Both BGP routers must support the route refresh capability (Cisco IOS 12.1 and later releases).
configured inbound soft reset (uses the **neighbor soft-reconfiguration** command)	Can be used when both BGP routers do not support the automatic route refresh capability.	Requires preconfiguration. Stores all received (inbound) routing policy updates without modification; is memory-intensive. Recommended only when absolutely necessary.

 Copyright © 2003, Cisco Systems, Inc.

Example


Route Refresh (Cont.)
Cisco.com
RTR-A
BGP session
RTR-B
Route refresh message
BGP routes are resent
Router#clear ip bgp neighbor neighbor in
Step #1 - Route refresh is negotiated when the BGP session is established
Step #2 - Inbound routing policy is changed on RTR-B
Step #3 - Operator requests inbound route refresh
Step #4 - RTR-B sends route refresh message to RTR-A
Step #5 - RTR-A resends all BGP routes to RTR-B
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-19

The capability to use the route refresh feature must be negotiated by the router when the BGP session is first established. The local router keeps a record that the capability is available with the neighbor. There is no need to keep a copy of the routing information received from the neighbor if it has the capability to refresh.

After reconfiguring the filters and route-maps that will implement a new routing policy, a network administrator can issue the **clear ip bgp** *ip-address* **soft in** command in the local router. The router will check whether the route refresh capability is available, and if it is, requests a resend of the BGP table of the neighbor instead of replaying its own copy.

Practice

Q1) What are two situations where you would you prefer inbound soft reconfiguration to route refresh? (Choose two.)

A) when there is insufficient memory to hold a copy of the BGP table of the neighbor

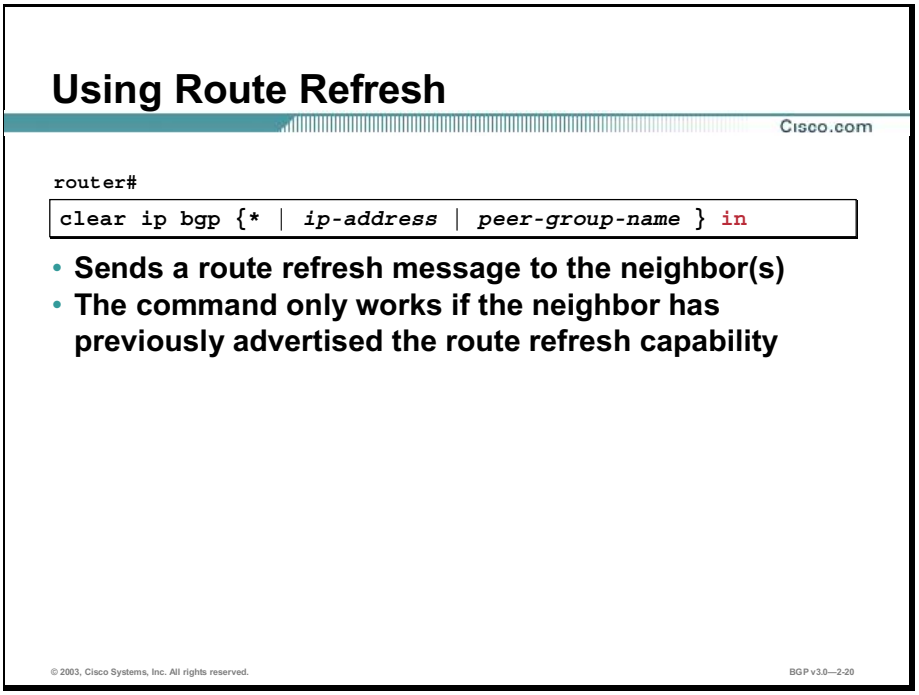
B) when a route refresh fails

C) when you wish to troubleshoot filters and use the **show ip bgp neighbor** command with the **received-routes** option

D) when the neighboring router does not support the route refresh capability

Using Route Refresh

This topic identifies the Cisco IOS commands required to perform a route refresh.


Using Route Refresh
Cisco.com
router#
clear ip bgp {* | ip-address | peer-group-name } in
• Sends a route refresh message to the neighbor(s)
• The command only works if the neighbor has previously advertised the route refresh capability
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-20

Use the **clear ip bgp * in** command to send a route refresh message to all neighbors or **clear ip bgp** *ip-address* **in** to send a route refresh message to a specific neighbor.

You need not use the **soft** keyword, because soft reset is automatically assumed when the route refresh capability is supported.

clear ip bgp

To reset a BGP connection using BGP soft reconfiguration, use the **clear ip bgp** privileged EXEC command at the system prompt.

clear ip bgp {* | *ip-address* | *peer-group-name*} [**soft** [**in** | **out**]]

Syntax Description

*	Resets all current BGP sessions.
address	Resets only the identified BGP neighbor.
peer-group-name	Resets the specified BGP peer group.
soft	(Optional) Soft reset. Does not reset the session.
in | **out**	(Optional) Triggers inbound or outbound soft reconfiguration. If the **in** or **out** option is not specified as a command argument, both inbound and outbound soft reset are triggered. Uses a route refresh message if it is supported by the neighbor.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What command do you use to trigger route refresh to a specific BGP neighbor?

A) **clear ip bgp** *ip-address* **in**

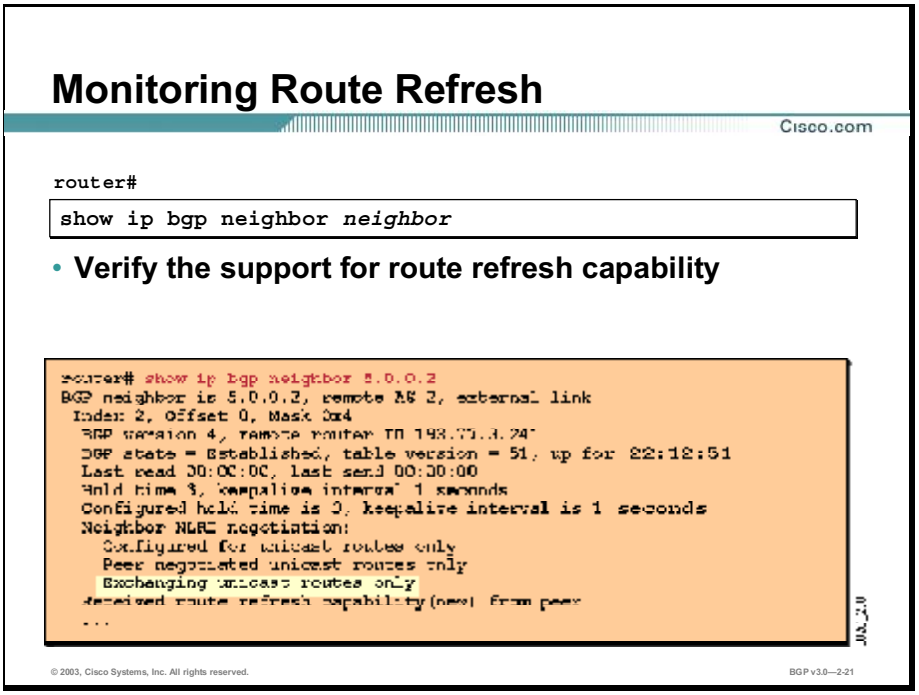
B) **show ip bgp neighbor**

C) **clear ip bgp ***

D) **clear ip bgp** *ip-address* **soft out**

Monitoring Route Refresh

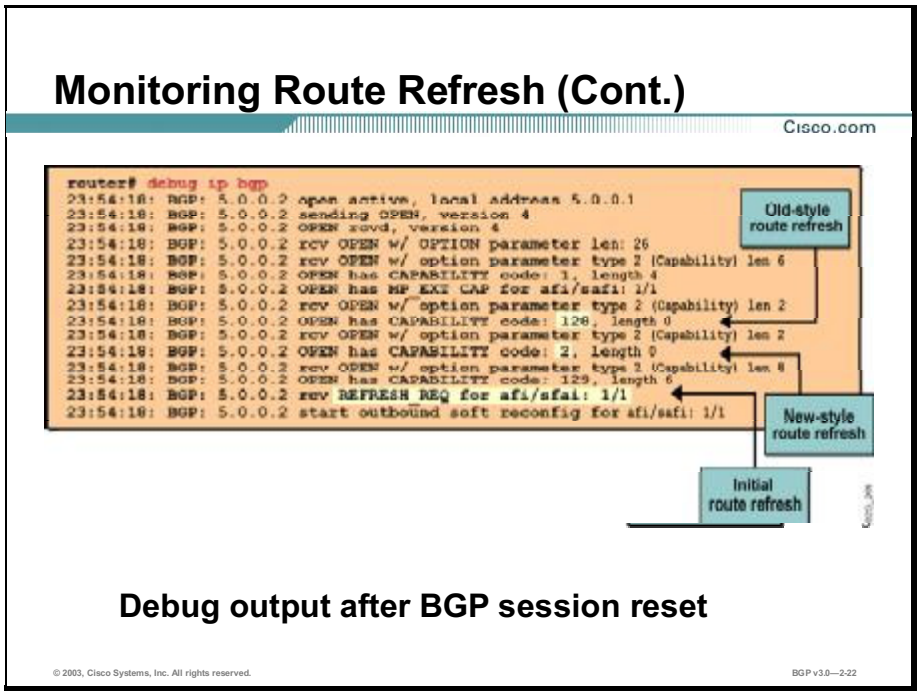
This topic identifies the Cisco IOS commands required to monitor route refresh operation.


Monitoring Route Refresh
Cisco.com
router#
show ip bgp neighbor neighbor
• Verify the support for route refresh capability
router# show ip bgp neighbor 5.0.0.2
BGP neighbor is 5.0.0.2, remote AS 2, external link
Index 2, Offset 0, Mask 0x4
BGP state = Established, table version = 51, up for 22:12:51
Last read 00:00:00, last send 00:00:00
Hold time 9, keepalive interval 1 seconds
Configured hold time is 9, keepalive interval is 1 seconds
Neighbor NLRI negotiation:
Configured for unicast routes only
Peer negotiated unicast routes only
Exchanging unicast routes only
...
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-21

Use the **show ip bgp neighbor** command to see if the neighbor supports the route refresh message.

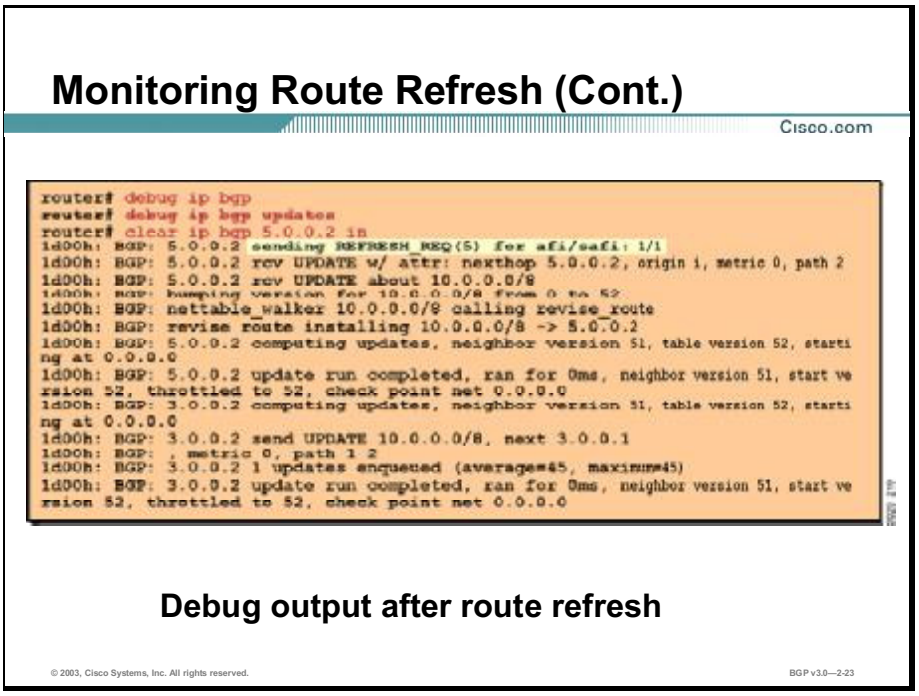
Note	The printout of the **show ip bgp neighbor** command varies between IOS releases. The printout in the figure here was generated by Cisco IOS Release 12.0(1)S and reflects the command output generated when manually configuring soft reset.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Route Refresh (Cont.)
Cisco.com
router# debug ip bgp
23:54:18: BGP: 5.0.0.2 open active, local address 5.0.0.1
23:54:18: BGP: 5.0.0.2 sending OPEN, version 4
23:54:18: BGP: 5.0.0.2 OPEN rcvd, version 4
23:54:18: BGP: 5.0.0.2 rcv OPEN w/ OPTION parameter len: 26
23:54:18: BGP: 5.0.0.2 rcv OPEN w/ option parameter type 2 (Capability) len 6
23:54:18: BGP: 5.0.0.2 OPEN has CAPABILITY code: 1, length 4
23:54:18: BGP: 5.0.0.2 OPEN has MP_EXT CAP for afi/safi: 1/1
23:54:18: BGP: 5.0.0.2 rcv OPEN w/ option parameter type 2 (Capability) len 2
23:54:18: BGP: 5.0.0.2 OPEN has CAPABILITY code: 128, length 0
23:54:18: BGP: 5.0.0.2 rcv OPEN w/ option parameter type 2 (Capability) len 2
23:54:18: BGP: 5.0.0.2 OPEN has CAPABILITY code: 2, length 0
23:54:18: BGP: 5.0.0.2 rcv OPEN w/ option parameter type 2 (Capability) len 0
23:54:18: BGP: 5.0.0.2 OPEN has CAPABILITY code: 129, length 6
23:54:18: BGP: 5.0.0.2 rcv REFRESH REQ for afi/safi: 1/1
23:54:18: BGP: 5.0.0.2 start outbound soft reconfig for afi/safi: 1/1
Old-style route refresh
New-style route refresh
Initial route refresh
Debug output after BGP session reset
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-22

Use **debug ip bgp** to display the negotiation of capabilities. Debugging displays received capabilities.

The example shows that a neighbor advertises both old-style and standard (new-style) route refresh. After the session has been established, an initial standard route refresh message was sent by the router for the address family 1/1 (IP version 4 [IPv4] unicast).


Monitoring Route Refresh (Cont.)
Cisco.com
router# debug ip bgp
router# debug ip bgp updates
router# clear ip bgp 5.0.0.2 in
1d00h: BGP: 5.0.0.2 sending REFRESH_REQ(5) for afi/safi: 1/1
1d00h: BGP: 5.0.0.2 rcv UPDATE w/ attr: nexthop 5.0.0.2, origin i, metric 0, path 2
1d00h: BGP: 5.0.0.2 rcv UPDATE about 10.0.0.0/8
1d00h: BGP: nettable_walker 10.0.0.0/8 calling revise_route
1d00h: BGP: revise route installing 10.0.0.0/8 -> 5.0.0.2
1d00h: BGP: 5.0.0.2 computing updates, neighbor version 51, table version 52, starting at 0.0.0.0
1d00h: BGP: 5.0.0.2 update run completed, ran for 0ms, neighbor version 51, start version 52, throttled to 52, check point net 0.0.0.0
1d00h: BGP: 3.0.0.2 computing updates, neighbor version 51, table version 52, starting at 0.0.0.0
1d00h: BGP: 3.0.0.2 send UPDATE 10.0.0.0/8, next 3.0.0.1
1d00h: BGP: , metric 0, path 1 2
1d00h: BGP: 3.0.0.2 1 updates enqueued (average=45, maximum=45)
1d00h: BGP: 3.0.0.2 update run completed, ran for 0ms, neighbor version 51, start version 52, throttled to 52, check point net 0.0.0.0
Debug output after route refresh
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-23

Debugging also shows a route refresh message being sent to a neighbor after the network administrator issues the **clear ip bgp neighbor in** command from privileged Exec mode.

Practice

Q1) How do you determine whether a BGP neighbor supports route refresh?

A) A flag in the BGP table indicates the presence of route refresh capability.

B) The **show ip bgp neighbor** command indicates if the option is supported.

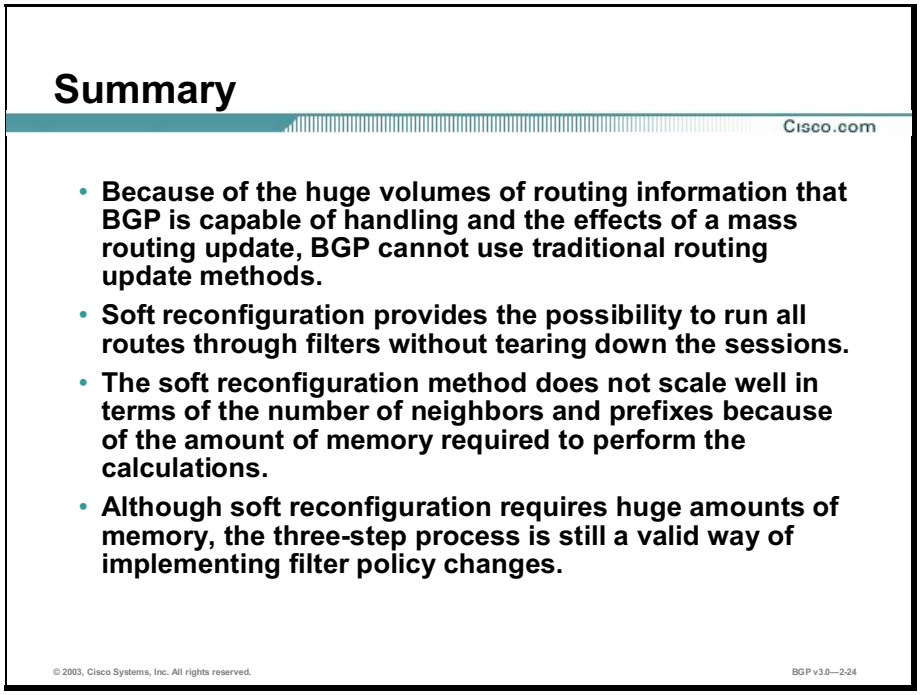
C) Initiate **debug ip bgp negotiation** to see if the router has completed a route refresh capabilities exchange.

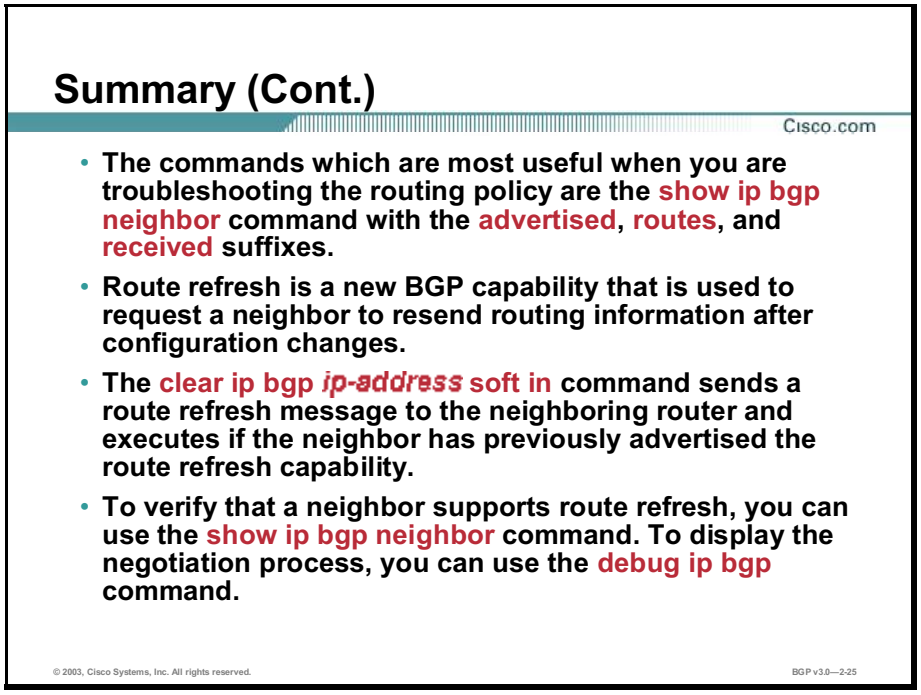
D) Execute the **clear ip bgp *** command. Command-line BGP status messages will indicate route refresh support capabilities.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Because of the huge volumes of routing information that BGP is capable of handling and the effects of a mass routing update, BGP cannot use traditional routing update methods.
• Soft reconfiguration provides the possibility to run all routes through filters without tearing down the sessions.
• The soft reconfiguration method does not scale well in terms of the number of neighbors and prefixes because of the amount of memory required to perform the calculations.
• Although soft reconfiguration requires huge amounts of memory, the three-step process is still a valid way of implementing filter policy changes.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-24


Summary (Cont.)
Cisco.com
• The commands which are most useful when you are troubleshooting the routing policy are the show ip bgp neighbor command with the advertised, routes, and received suffixes.
• Route refresh is a new BGP capability that is used to request a neighbor to resend routing information after configuration changes.
• The clear ip bgp ip-address soft in command sends a route refresh message to the neighboring router and executes if the neighbor has previously advertised the route refresh capability.
• To verify that a neighbor supports route refresh, you can use the show ip bgp neighbor command. To display the negotiation process, you can use the debug ip bgp command.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—2-25

Next Steps

After completing this lesson, go to:

- Route Selection Using Attributes module

References

For additional information, refer to these resources:

- For more information on route refresh and soft reconfiguration, refer to “BGP Soft Reset Enhancement” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t7/sftrst.htm

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: Implementing Changes in BGP Policy

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure the soft reconfiguration feature to minimize the impact of expediting BGP policy updates, given a network scenario where you must support connections to multiple ISPs.

After completing this exercise, you will be able to:

- Configure and perform a soft reconfiguration
- Monitor the operations of a soft reconfiguration

Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor *ip-address* **soft-reconfiguration inbound**	Enable inbound soft reconfiguration.
show ip bgp	Inspect the contents of the BGP table.
show ip neighbor *ip-address*	Show detailed information about a specific neighbor.
show ip neighbor *ip-address* **received-routes**	View all updates received from the specified neighbor (before filters).
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- Your customers complain that they are sometimes unable to reach your e-commerce servers. You suspect that the frequent changes in your BGP routing policies might be the reason and you would like to implement a mechanism that enables you to perform nondisruptive changes in routing policies. Unfortunately, your service providers run a Cisco IOS release that does not support BGP route refresh functionality.

- In this exercise, you will configure soft reconfiguration to enable nondisruptive changes in BGP routing policies.

- Figure 1 displays the required BGP connectivity within your workgroup as well as the BGP sessions with two different service providers.

 Copyright © 2003, Cisco Systems, Inc.

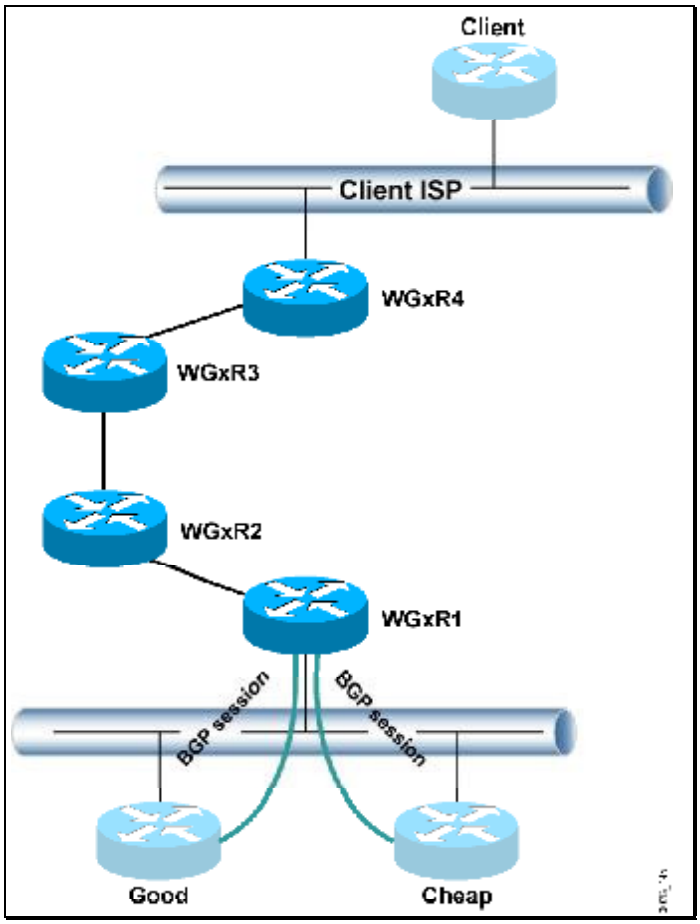


Figure 1: Connecting to two different service providers

Exercise Procedure

Complete these steps:

Step 1 Before you begin, check to ensure that soft reconfiguration is not configured.

```
wg1r1#sh ip bgp ne 192.168.20.20 received
% Inbound soft reconfiguration not enabled

wg1r1#sh ip bgp ne 192.168.20.22 received
% Inbound soft reconfiguration not enabled
```

Step 2 Enable inbound soft reconfiguration on WGxR1 to both neighbors “Good” and “Cheap.”

Note	Using the **neighbor soft-reconfiguration** command may require your router to have much more memory than normally needed.

Exercise Verification

You have completed this exercise when you attain these results:

```
wg1r1#sh ip bgp ne 192.168.20.20 received

BGP table version is 50, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.0.0.0         192.168.20.233           0             0 20 i
*  128.20.0.0       192.168.20.20            0             0 20 i
*> 128.20.12.0/24   192.168.20.20            0             0 20 i
*  128.22.0.0       192.168.20.20            0             0 20 i
*> 128.22.12.0/24   192.168.20.22                          0 20 22 i
*  128.26.0.0       192.168.20.20            0             0 20 42 26
i
*  128.37.0.0       192.168.20.20            0             0 20 42 37
i
*  128.42.0.0       192.168.20.20            0             0 20 42 i
*  128.51.0.0       192.168.20.20            0             0 20 42 26
51 i
*> 128.213.0.0      192.168.20.20            0             0 20 213 i
*  128.214.0.0      192.168.20.22                          0 20 22
214 i
*  192.20.11.0      192.168.20.20            0             0 20 i
*  192.20.12.0/30   192.168.20.20            0             0 20 i
*  192.22.11.0      192.168.20.20            0             0 20 i
*  192.22.12.0/30   192.168.20.22                          0 20 22 i
*  192.26.11.0      192.168.20.20            0             0 20 42 26
i
*  192.37.11.0      192.168.20.20            0             0 20 42 37
i
*  192.42.11.0      192.168.20.20            0             0 20 42 i
*  192.51.11.0      192.168.20.20            0             0 20 42 26
51 i
*  192.213.11.0     192.168.20.20            0             0 20 213 i
*  192.214.11.0     192.168.20.22                          0 20 22
214 i
*  200.20.0.0/16    192.168.20.20            0             0 20 i
*  200.22.0.0/16    192.168.20.22                          0 20 22 i

Total number of prefixes 23
```

```
wg1r1#sh ip bgp ne 192.168.20.22 received
BGP table version is 50, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.0.0.0         192.168.20.233                         0 22 20 i
*> 128.20.0.0       192.168.20.22            0           100 22 i
*  128.20.12.0/24   192.168.20.20                          0 22 20 i
*> 128.22.0.0       192.168.20.22            0           100 22 i
*  128.22.12.0/24   192.168.20.22            0             0 22 i
*> 128.26.0.0       192.168.20.22            0           100 22 26 i
*> 128.37.0.0       192.168.20.22            0           100 22 26 42
37 i
*> 128.42.0.0       192.168.20.22            0           100 22 26 42
i
```

 Copyright © 2003, Cisco Systems, Inc.

```
*> 128.51.0.0       192.168.20.22           0           100 22 26 51
i
*  128.213.0.0      192.168.20.20                         0 22 20
213 i
*> 128.214.0.0      192.168.20.22           0           100 22 214 i
*> 192.20.11.0      192.168.20.22           0           100 22 i
*> 192.20.12.0/30   192.168.20.20                       100 22 20 i
*> 192.22.11.0      192.168.20.22           0           100 22 i
*> 192.22.12.0/30   192.168.20.22           0           100 22 i
*> 192.26.11.0      192.168.20.22           0           100 22 26 i
*> 192.37.11.0      192.168.20.22           0           100 22 26 42
37 i
*> 192.42.11.0      192.168.20.22           0           100 22 26 42
i
*> 192.51.11.0      192.168.20.22           0           100 22 26 51
i
*  192.213.11.0     192.168.20.20                         0 22 20
213 i
*> 192.214.11.0     192.168.20.22           0           100 22 214 i
*> 200.20.0.0/16    192.168.20.20                       100 22 20 i
*> 200.22.0.0/16    192.168.20.22           0           100 22 i

Total number of prefixes 23
```

Answer these questions:

Q1) What command do you use to show which entries in the BGP table of the local router have been propagated to a specific neighbor?

Q2) What command do you use to display which of the routes in the local BGP table have been received and accepted from an indicated neighbor?

 Copyright © 2003, Cisco Systems, Inc.

3

Route Selection Using Attributes

Overview

Routes learned via the Border Gateway Protocol (BGP) have properties associated with them that aid a router in determining the best route to a destination when multiple paths to that particular destination exist. These properties are referred to as BGP attributes. This module introduces the role of BGP attributes, and how their presence influences route selection in BGP. Understanding how BGP attributes influence route selection is required for the design of robust networks.

This module provides advanced information on how to connect Internet customers to multiple service providers. It includes an in-depth description of BGP attributes used in route selection, including weight, local preference, AS-path prepending, multi-exit discriminator (MED), and BGP communities.

Upon completing this module, you will be able to:

- List BGP path attributes and the functionality of each attribute.
- Successfully configure BGP to influence route selection using the weight attribute, given a customer scenario where you must support multiple connections.
- Use the local preference attribute to influence route selection, given a customer scenario where you must support multiple connections.
- Use AS-path prepending to influence the return path selected by the neighboring autonomous systems, given a customer scenario where you must support multiple connections.

- Use the MED attribute to influence route selection, given a customer scenario where you must support multiple connections.
- Use BGP community attributes to influence route selection, given a customer scenario where you must support multiple connections.

Outline

The module contains these lessons:

- BGP Path Attributes
- Influencing BGP Route Selection with Weights
- BGP Local Preference
- AS-Path Prepending
- BGP Multi-Exit Discriminator
- BGP Communities

 Copyright © 2003, Cisco Systems, Inc.

BGP Path Attributes

Overview

This lesson introduces Border Gateway Protocol (BGP) attributes and their purpose. The lesson also discusses classifications used to describe attributes and the properties of each classification. The functionality of the AS-path and next-hop attributes are also explained in detail in this lesson.

Importance

To aid routers in calculating the best route to select when multiple paths to a particular destination exist, routes learned via BGP have properties associated with them. These properties are referred to as BGP attributes, and an understanding of how BGP attributes influence route selection is required to design robust BGP networks.

Objectives

Upon completing this lesson, you will be able to:

- Describe the purpose of BGP path attributes
- Explain the difference between mandatory and discretionary well-known BGP attributes
- Explain the difference between nontransitive and transitive optional BGP attributes
- Describe the functionality of the AS-path attribute
- Describe the functionality of the next-hop attribute

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

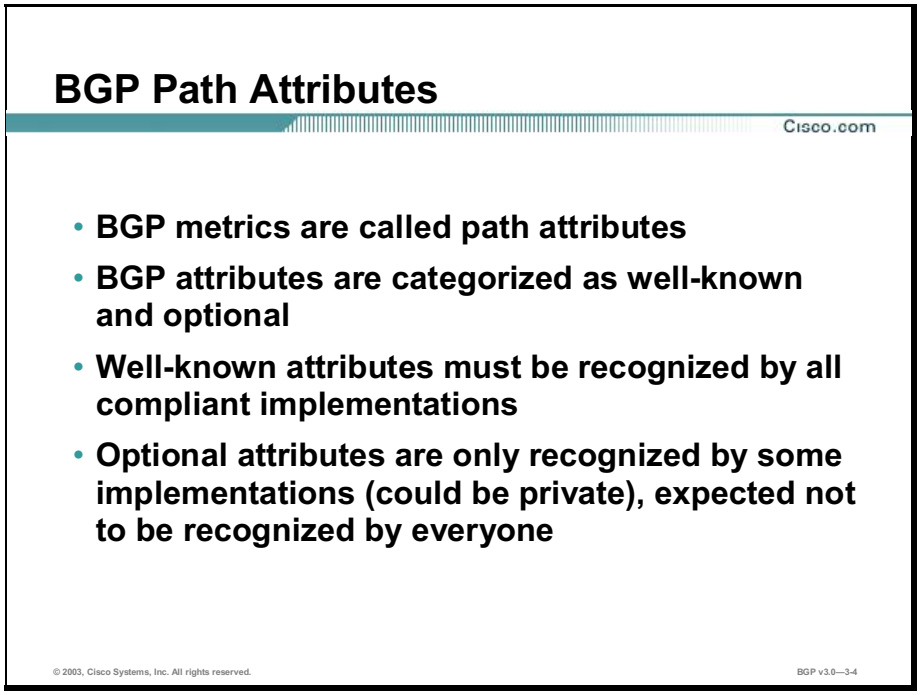
This lesson includes these topics:

- Overview
- BGP Path Attributes
- Well-Known BGP Attributes
- Optional BGP Attributes
- AS-Path Attribute
- Next-Hop Attribute
- Summary
- Assessment (Quiz): BGP Path Attributes

 Copyright © 2003, Cisco Systems, Inc.

BGP Path Attributes

This topic describes the concept of BGP path attributes.


BGP Path Attributes
Cisco.com
• BGP metrics are called path attributes
• BGP attributes are categorized as well-known and optional
• Well-known attributes must be recognized by all compliant implementations
• Optional attributes are only recognized by some implementations (could be private), expected not to be recognized by everyone
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-4

Each BGP update consists of one or more IP subnets and a set of attributes attached to them. Some of the attributes are required to be recognized by all BGP implementations. Those attributes are called “well-known BGP attributes.”

Attributes that are not well-known are called “optional.” These could be attributes specified in a later extension of BGP or even private vendor extensions not documented in a standard document.

Practice

Q1) What do network administrators use in BGP to define the metrics used for best route selection?

A) link states

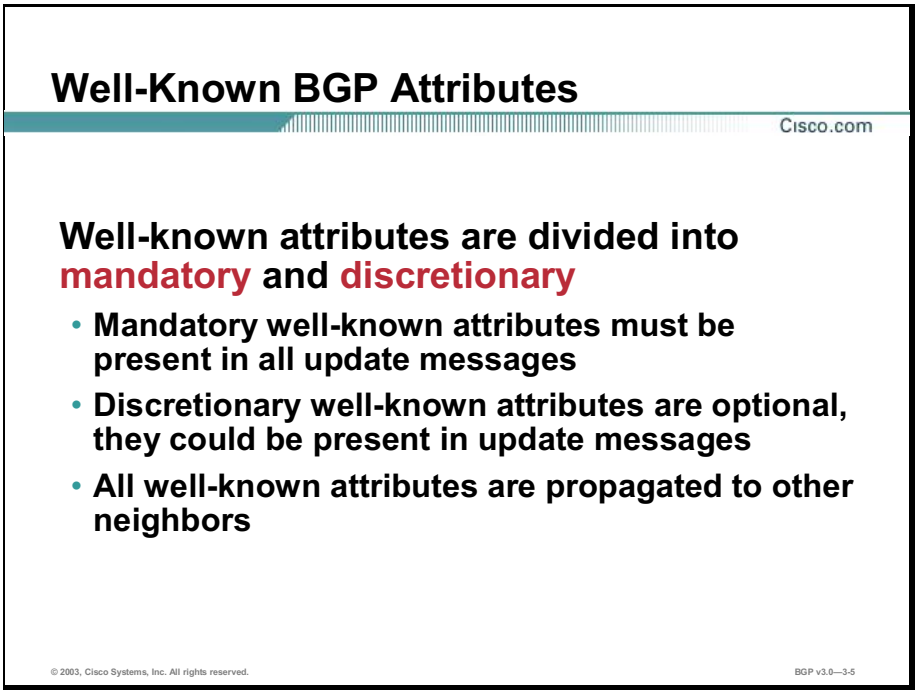
B) path attributes

C) distance vectors

D) cost

Well-Known BGP Attributes

This topic explains the differences between the well-known BGP mandatory and discretionary attributes.

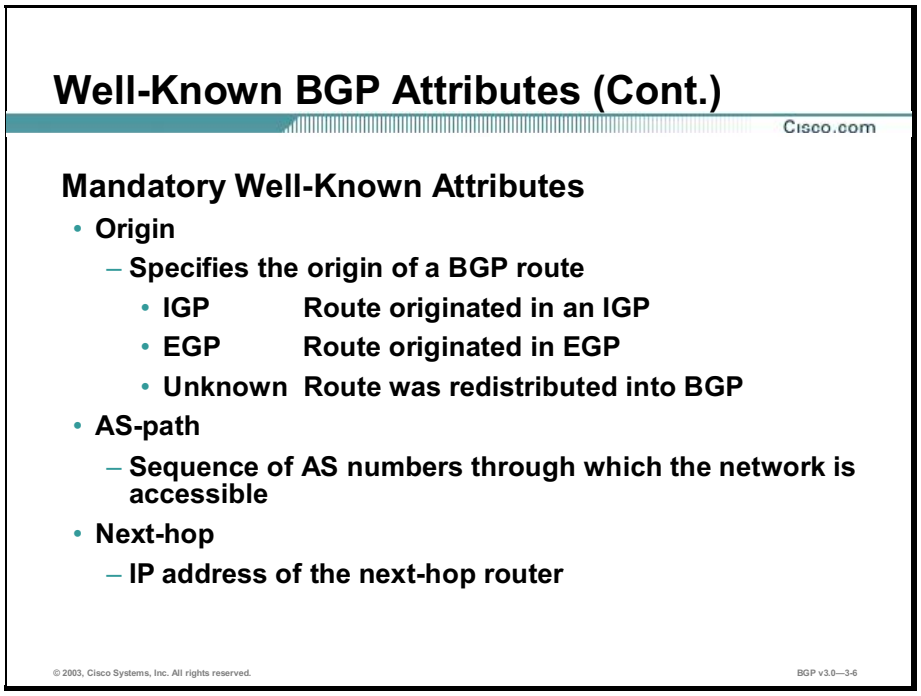

Well-Known BGP Attributes
Cisco.com
Well-known attributes are divided into mandatory and discretionary
• Mandatory well-known attributes must be present in all update messages
• Discretionary well-known attributes are optional, they could be present in update messages
• All well-known attributes are propagated to other neighbors
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-5

There is a small set of three specific well-known attributes that are required to be present on every update. These are the next-hop, AS-path, and origin attributes and are referred to as “mandatory well-known attributes.”

Other well-known attributes may or may not be present depending on the circumstances under which the updates are sent and the desired routing policy. The well-known attributes that could be present, but are not required, are called “discretionary well-known attributes.”

When a router receives a BGP update, it will analyze the attached attributes and compare them with the attributes attached to the same IP subnet when received from a different source. The router then makes a decision about which source indicates the best path to the particular IP subnet. The best route is propagated, along with its well-known attributes, to other BGP-speaking neighbors.

 Copyright © 2003, Cisco Systems, Inc.

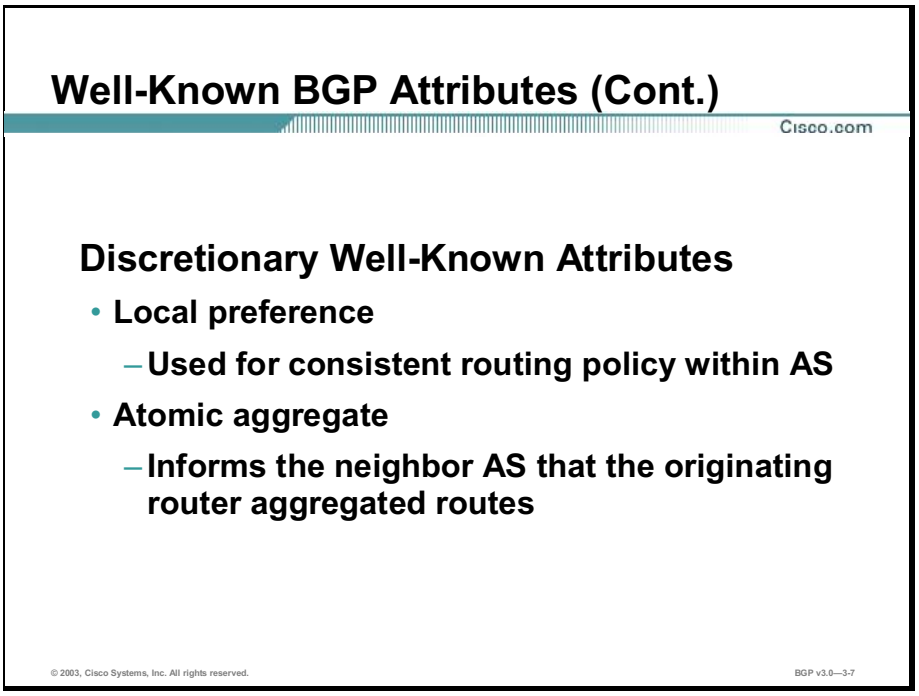

Well-Known BGP Attributes (Cont.)
Cisco.com
Mandatory Well-Known Attributes
• Origin
– Specifies the origin of a BGP route
• IGP Route originated in an IGP
• EGP Route originated in EGP
• Unknown Route was redistributed into BGP
• AS-path
– Sequence of AS numbers through which the network is accessible
• Next-hop
– IP address of the next-hop router
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-6

The three mandatory well-known attributes are origin, AS-path, and next-hop.

- When a router first originates a route in BGP, it sets the origin attribute. If information about an IP subnet is injected using the **network** command or via aggregation (route summarization within BGP), the origin attribute is set to "IGP." If information about an IP subnet is injected using redistribution, the origin attribute is set to "unknown" or "incomplete" (these two words have the same meaning). The origin code, "EGP," was used when the Internet was migrating from exterior gateway protocol (EGP) to BGP and is now obsolete.

- The egress router modifies the AS-path attribute every time information about a particular IP subnet passes over an autonomous system (AS) border. When a router first originates a route in BGP, the AS-path attribute is empty. Each time that the route crosses an AS boundary, the transmitting AS prepends its own AS number to appear first in the AS path. You can track the sequence of autonomous systems through which the route has passed by using the AS-path attribute.

- The router also modifies the next-hop attribute as the route passes through the network. It indicates the IP address of the next-hop router—the router to which the receiving router should forward the IP packets toward the destination advertised in the routing update.


Well-Known BGP Attributes (Cont.)
Cisco.com
Discretionary Well-Known Attributes
• Local preference
– Used for consistent routing policy within AS
• Atomic aggregate
– Informs the neighbor AS that the originating router aggregated routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-7

Discretionary well-known attributes must be supported by all BGP implementations but do not have to be present in all BGP updates. Routers use discretionary well-known attributes only when their functions are required.

- Local preference is used in the route selection process. This attribute is carried within an AS only. The router prefers a route with a high local preference value to a route with a low value. By default, routes received from a peer AS are tagged with the local preference set to a value of 100 before they are entered into the local AS. If this value is changed through BGP configuration, the BGP selection process is influenced. Because all routers within the AS get the attribute along with the route, a consistent routing decision is made throughout the AS.

- The atomic aggregate attribute is attached to a route that is created as a result of route summarization (called aggregation in BGP). It signals that information that was present in the original routing updates may have been lost when the updates were summarized into a single entry.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Which three BGP path attributes must be carried with each update? (Choose three.)

A) origin

B) AS-path

C) local preference

D) next-hop

Q2) Which two attributes are discretionary well-known BGP attributes? (Choose two.)

A) multi-exit discriminator

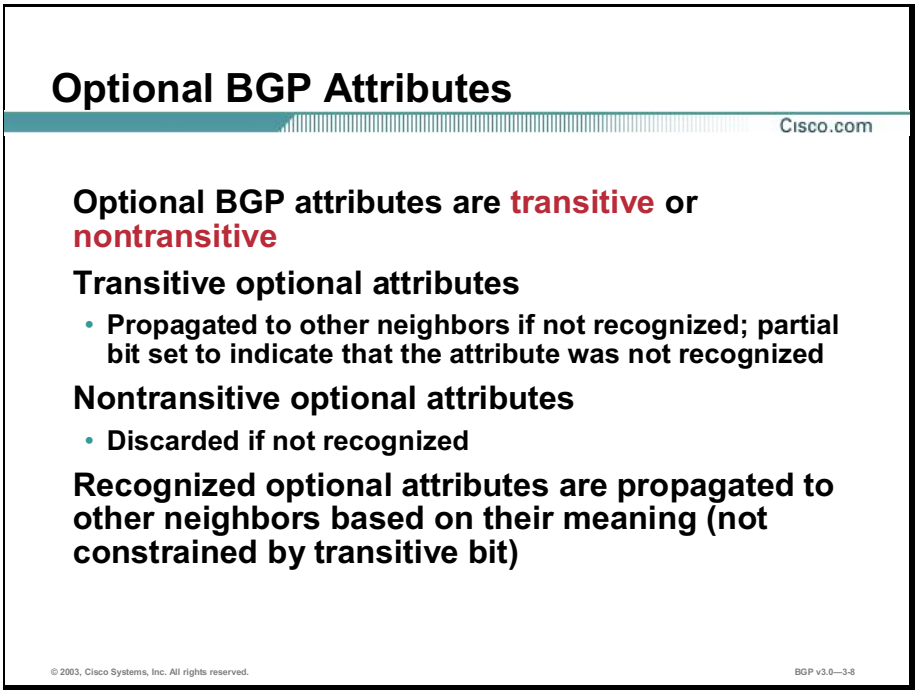
B) local preference

C) origin

D) atomic aggregate

Optional BGP Attributes

This topic explains the difference between the transitive and nontransitive optional BGP attributes.

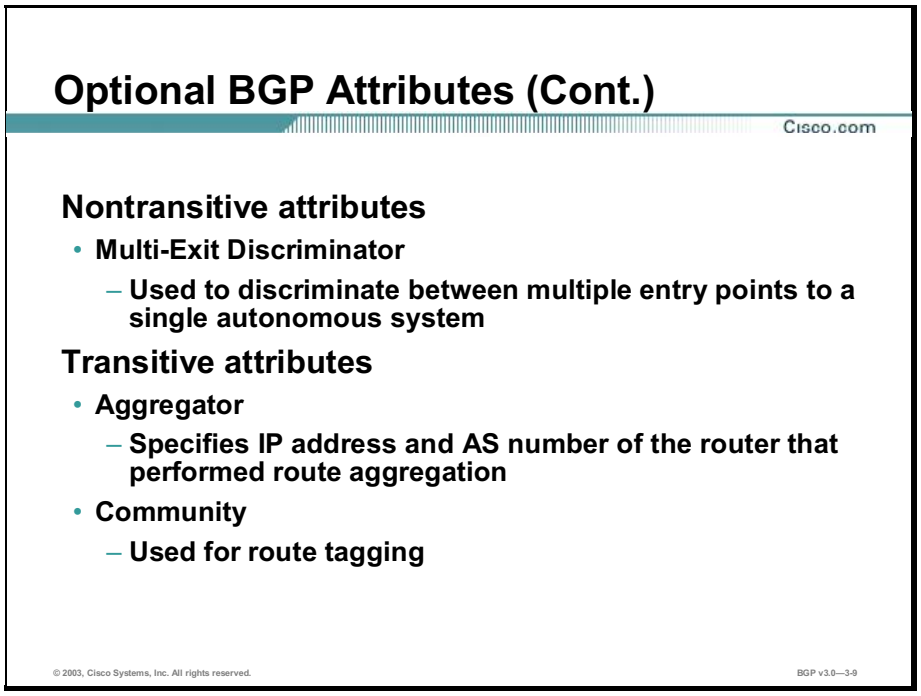

Optional BGP Attributes
Cisco.com
Optional BGP attributes are transitive or nontransitive
Transitive optional attributes
• Propagated to other neighbors if not recognized; partial bit set to indicate that the attribute was not recognized
Nontransitive optional attributes
• Discarded if not recognized
Recognized optional attributes are propagated to other neighbors based on their meaning (not constrained by transitive bit)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-8

When a router receives an update that contains an optional attribute, the router checks if its implementation recognizes the particular attribute. If it does, then the router should know how to handle it and whether to propagate it or not.

If the router does not recognize the attribute, the BGP implementation should look for the transitive bit in the attribute code. Some attributes, although not recognized by the router, might still be helpful to upstream routers and should be propagated. These attributes (called “transitive optional attributes”) are propagated even when they are not recognized. If a router propagates an unknown transitive optional attribute, it will set an additional bit in the attribute header, called the “partial bit,” to indicate that at least one of the routers in the path did not recognize the meaning of a transitive optional attribute.

Other attributes, called “nontransitive optional attributes,” might be of no value to upstream routers if some router in the path does not recognize them. Routers that do not recognize these attributes will drop them.

 Copyright © 2003, Cisco Systems, Inc.


Optional BGP Attributes (Cont.)
Cisco.com
Nontransitive attributes
• Multi-Exit Discriminator
– Used to discriminate between multiple entry points to a single autonomous system
Transitive attributes
• Aggregator
– Specifies IP address and AS number of the router that performed route aggregation
• Community
– Used for route tagging
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-9

One of the nontransitive optional attributes is the multi-exit discriminator (MED) attribute, which also influences the BGP route selection process. Whenever there are several links between two adjacent autonomous systems, one AS can use the MED attribute to tell another AS to prefer one of the links for specific destinations.

Transitive optional attributes include:

- **Aggregator:** Identifies the AS and the router within that AS that created a route summarization, or aggregate.

- **Community:** A numerical value that can be attached to certain routes as they pass a specific point in the network. The community value can later be examined by other routers at different points in the network for filtering or route selection purposes. BGP configuration may cause routes with a specific community value to be treated differently than others.

Practice

Q1) How are recognized transitive optional attributes propagated between BGP neighbors?

A) Optional attributes are converted to transitive well-known attributes.

B) With the partial bit set.

C) Based on their meaning.

D) Optional attributes are not propagated to neighbors.

Q2) How are nonrecognized transitive optional attributes propagated between BGP neighbors?

A) Optional attributes are converted to transitive well-known attributes.

B) With the partial bit set.

C) Based on their meaning.

D) Optional attributes are not propagated to neighbors.

Q3) Which is a nontransitive optional BGP path attribute?

A) local preference

B) weight

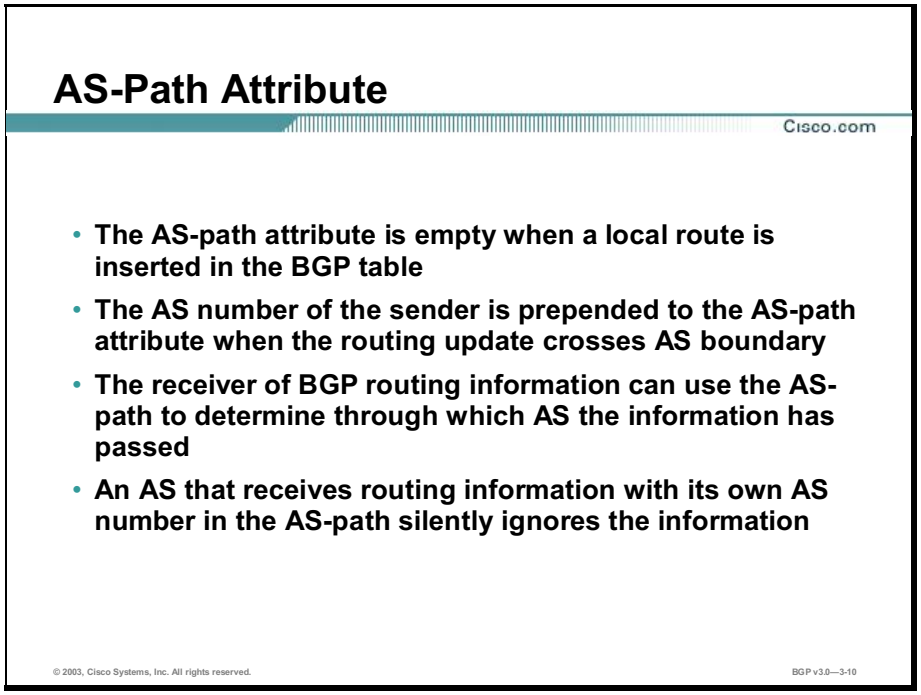
C) MED

D) community

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Attribute

This topic describes the functionality of the BGP AS-path attribute.


AS-Path Attribute
Cisco.com
• The AS-path attribute is empty when a local route is inserted in the BGP table
• The AS number of the sender is prepended to the AS-path attribute when the routing update crosses AS boundary
• The receiver of BGP routing information can use the AS-path to determine through which AS the information has passed
• An AS that receives routing information with its own AS number in the AS-path silently ignores the information
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-10

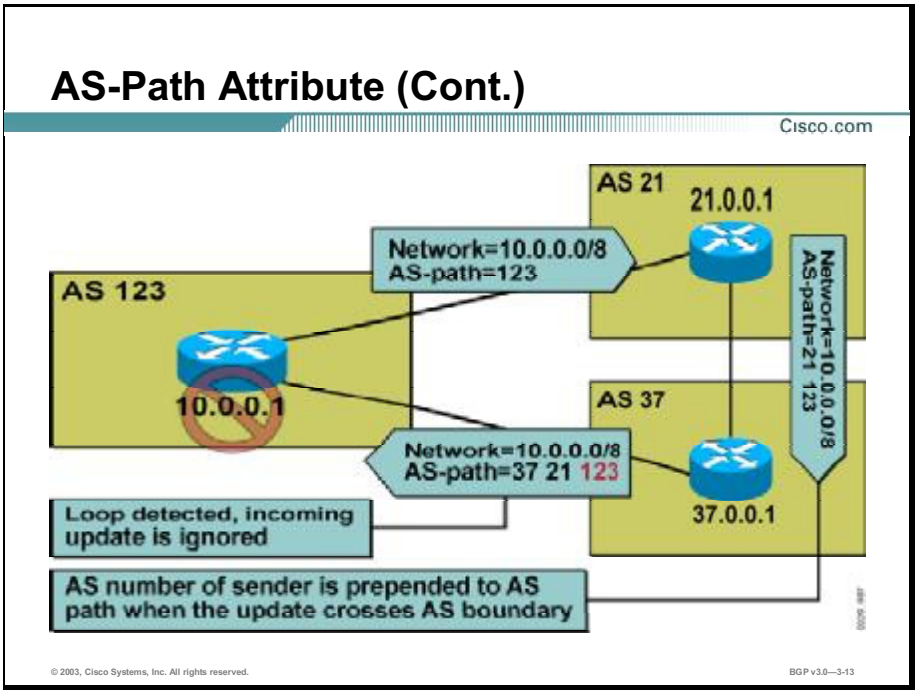
The AS-path attribute is modified by an edge router every time information about a particular IP subnet passes over an AS border. When a router first originates a route in BGP, the AS-path attribute is empty. The local AS number is prepended to the AS path each time the route crosses an AS boundary. There are several consequences of this behavior:

- When you examine BGP routes, the AS path can be interpreted as the sequence of autonomous systems that must be passed through in order to reach the indicated network. The AS that originally injected the route into BGP is always found in the rightmost end of the AS path.

- It is easy to distinguish local routes from routes received from other autonomous systems—BGP routes with an empty AS path were injected into BGP from within the local AS.

The AS-path attribute is also used to avoid routing loops. When a router receives a BGP update, it will check the AS-path attribute and look for its own AS number. If it is found in the AS path, then the route has already crossed the local AS and the router is now faced with a routing information loop. To avoid this situation, the route is silently ignored.

Example



The figure shows how BGP loop prevention works.

The network 10.0.0.0/8 is local to AS 123. The router in AS 123 injects the route 10.0.0.0/8 into BGP with an empty AS-path attribute.

When the routing update about network 10.0.0.0/8 is sent by the edge router in AS 123 to AS 21, AS number 123 is prepended to the empty AS path, resulting in an AS path consisting of only 123. The sending router does the prepending as part of the outgoing BGP update processing. While the route is still within AS 123, the AS-path entry for AS 123 will not appear in the AS path.

The router in AS 21 propagates the information about the network 10.0.0.0/8 to AS 37. As it is sending the BGP update to AS 37, it prepends its own AS number to the AS path, resulting in an AS path consisting of the sequence of 21 123.

AS 37 also propagates the received route to AS 123. To avoid a routing loop, where AS 123 might try to reach its own network (10.0.0.0/8) via AS 37, BGP has a built-in mechanism where the router in AS 123 drops the incoming update as soon as it finds its own AS (123), in the AS path. No error will be signaled, because nothing is really wrong. It is merely the procedure used by BGP to avoid a routing information loop.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) How do BGP routers detect routing loops?

A) BGP routers check for the longest AS path in all routing updates.

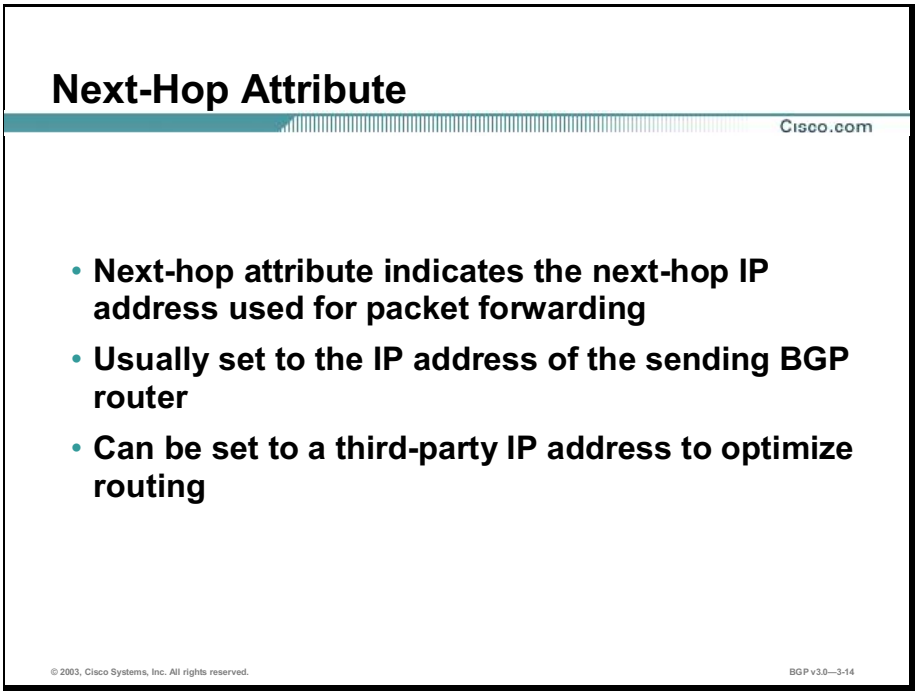
B) BGP routers ignore incoming routes with an AS path containing their own AS number.

C) Multiple paths with the same origin AS are considered routing loops.

D) Multiple paths with the same destination AS are considered routing loops.

Next-Hop Attribute

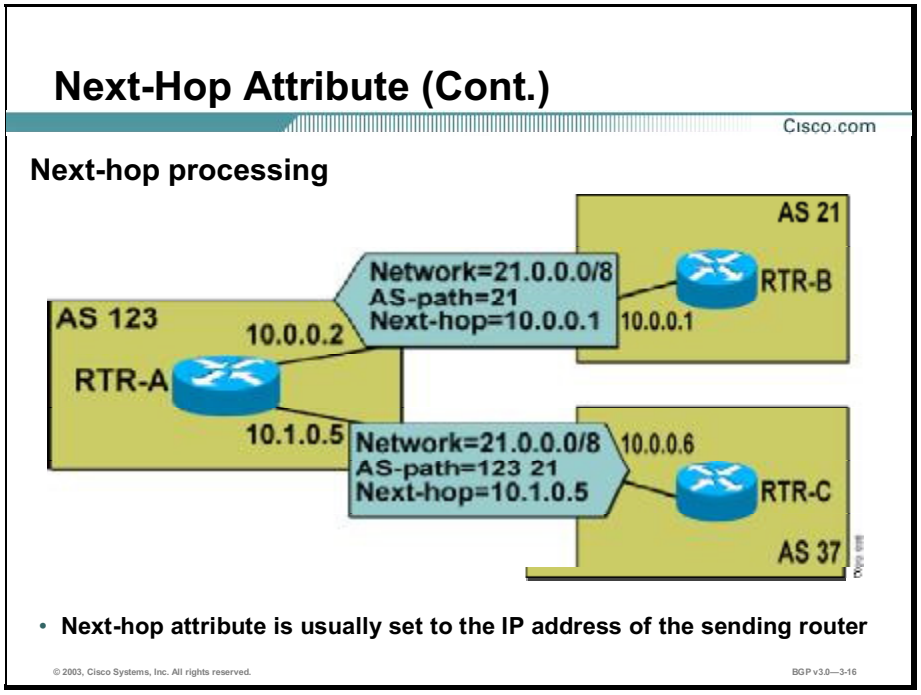
This topic describes the functionality of the next-hop attribute in BGP.


Next-Hop Attribute
Cisco.com
• Next-hop attribute indicates the next-hop IP address used for packet forwarding
• Usually set to the IP address of the sending BGP router
• Can be set to a third-party IP address to optimize routing
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-14

The BGP next-hop attribute identifies the IP address that a router should use to forward packets toward the destination announced in a BGP routing update. In most cases, the sending router sets the next-hop attribute to its own IP address. There are cases, however, where the next-hop IP address points to a third router.

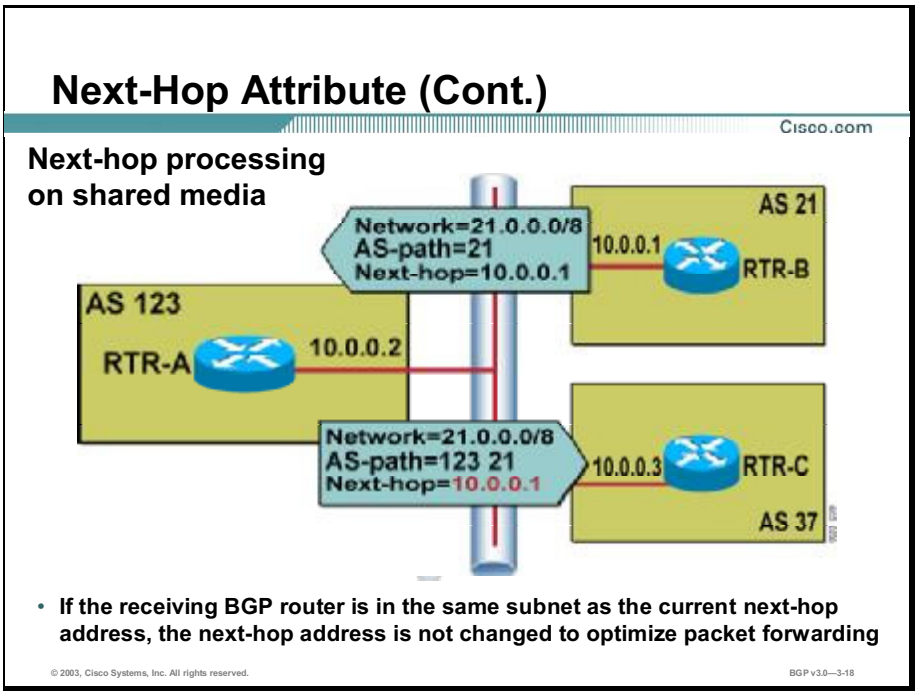
 Copyright © 2003, Cisco Systems, Inc.

Example



The figure shows the usual next-hop processing:

- RTR-B announces network 21.0.0.0/8 to RTR-A. The outgoing IP address of RTR-B (the address used to establish BGP TCP session) is used as the BGP next hop.
- RTR-A receives the routing update and installs it in its BGP table and routing table. Should RTR-A need to forward packets toward network 21.0.0.0/8, it would send those packets toward the IP address 10.0.0.1 (RTR-B)
- When RTR-A propagates the information about 21.0.0.0/8 to RTR-C, it sets the BGP next-hop attribute to its own IP address.



The next-hop processing changes if the BGP routers connect to a shared subnet. In the figure here, if RTR-A announces the network 21.0.0.0/8 to RTR-C with the BGP next-hop address set to RTR-A, the packets from AS 37 toward network 21.0.0.0/8 will have to cross the shared LAN twice. RTR-A thus sends the routing update toward RTR-C with the BGP next-hop address unchanged (still pointing toward RTR-B), allowing optimal data transfer across the shared LAN.

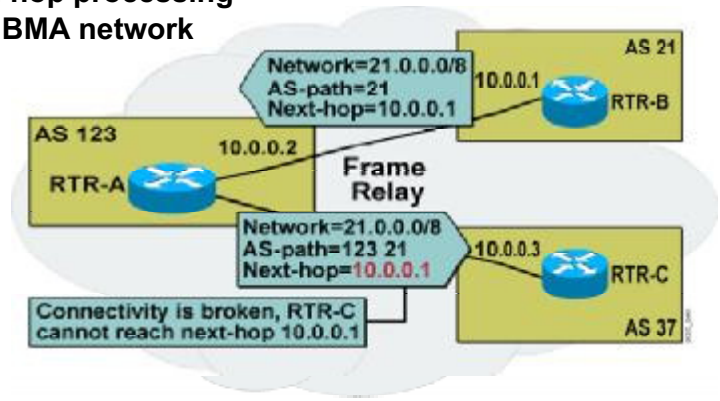
Note	More formally, the BGP next-hop rule states that if the current BGP next hop is in the same IP subnet as the receiving router, the next-hop address is not changed; otherwise, the next-hop attribute is changed to the IP address of the sending router.

 Copyright © 2003, Cisco Systems, Inc.

Next-Hop Attribute (Cont.)

Cisco.com

Next-hop processing on NBMA network



- **BGP next-hop processing can break connectivity with improper network designs over partially-meshed WAN networks**

© 2003, Cisco Systems, Inc. All rights reserved.

BGP next-hop processing results in optimum data transfer over shared media (for example, a LAN subnet). In partially meshed networks (like Frame Relay), BGP next-hop processing can break IP connectivity. Consider, for example, the network diagram above: RTR-A will send a routing update about network 21.0.0.0/8 to RTR-C with RTR-B set to the next-hop address (as they are all in the same subnet). Because there is no direct connection (virtual circuit) between RTR-C and RTR-B, but RTR-C still tries to send packets directly toward RTR-B, the connectivity between AS 37 and AS 21 is broken.

There are two ways to solve the connectivity loss introduced by this design:

- Use the subinterfaces on RTR-A to make sure that RTR-B and RTR-C are in different subnets (and BGP next-hop processing would ensure that RTR-A is the BGP next hop in the outgoing BGP updates).

- Disable the BGP next-hop processing on RTR-A. (This option is strongly discouraged in normal BGP designs because routing problems should be solved with a proper network design.)

Practice

Q1) When is the next-hop attribute different from the IP address of the sending router?

A) In situations where static routing is used to reach nondirectly connected IBGP peers.

B) In cases where the BGP next hop is in a different subnet than the receiving router.

C) The next-hop attribute is always different, because it points to the exit gateway.

D) If the current BGP next hop is in the same IP subnet as the receiving router.

Q2) How can improperly designed, partially meshed networks break BGP connectivity?

A) A direct connection may not be available to the same subnet next hop.

B) A design using separate IP networks and subinterfaces can cause BGP to set the incorrect next hop.

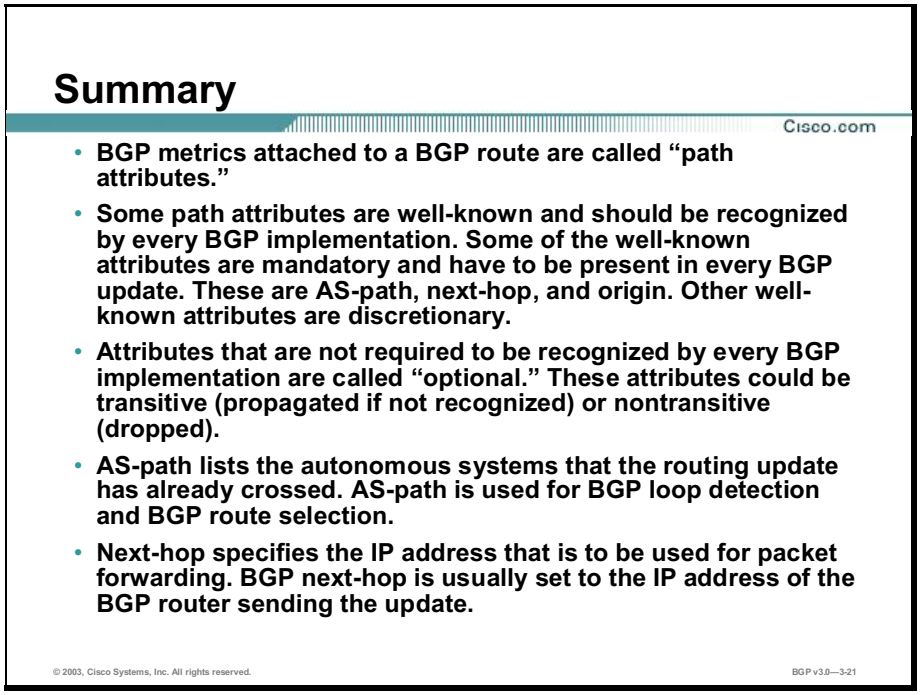
C) The BGP next hop is always set to the sending router, causing it to act as the hub of the meshed network.

D) Partially meshed networks automatically disable BGP next-hop processing and should be avoided.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• BGP metrics attached to a BGP route are called “path attributes.”
• Some path attributes are well-known and should be recognized by every BGP implementation. Some of the well-known attributes are mandatory and have to be present in every BGP update. These are AS-path, next-hop, and origin. Other well-known attributes are discretionary.
• Attributes that are not required to be recognized by every BGP implementation are called “optional.” These attributes could be transitive (propagated if not recognized) or nontransitive (dropped).
• AS-path lists the autonomous systems that the routing update has already crossed. AS-path is used for BGP loop detection and BGP route selection.
• Next-hop specifies the IP address that is to be used for packet forwarding. BGP next-hop is usually set to the IP address of the BGP router sending the update.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-21

Next Steps

After completing this lesson, go to:

- Influencing BGP Route Selection with Weights lesson

References

For additional information, refer to these resources:

- For more information on BGP attributes, refer to “Border Gateway Protocol” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

Quiz: BGP Path Attributes

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe the purpose of BGP path attributes
- Explain the difference between mandatory and discretionary well-known BGP attributes
- Explain the difference between nontransitive and transitive optional BGP attributes
- Describe the functionality of the AS-path attribute
- Describe the functionality of the next-hop attribute

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) Which three statements are true of BGP mandatory well-known attributes? (Choose three.)

A) They must be present in all BGP updates.

B) All BGP-compliant implementations must recognize them.

C) All BGP-compliant routers must adhere to policies specified in mandatory attributes.

D) All well-known attributes are propagated to other neighbors.

 Copyright © 2003, Cisco Systems, Inc.

Q2) Which three attributes are BGP mandatory well-known attributes? (Choose three.)

A) next-hop

B) weight

C) AS-path

D) origin

Q3) What three possible values are assigned to the BGP origin attribute? (Choose three.)

A) IGP

B) EGP

C) unknown

D) internal

Q4) What nontransitive optional BGP attribute is useful in assisting with the route selection process when multiple links to another AS exist?

A) next-hop

B) local preference

C) MED

D) AS-path

Q5) How is the BGP next-hop attribute modified?

A) If the next-hop attribute is in the same IP subnet as the receiving router, the attribute is unchanged; otherwise, it is set to the IP address of the sending router.

B) The next-hop attribute is always set to the IP address of the sending router.

C) The next-hop attribute is modified only when BGP packets exit an AS.

D) The BGP next-hop attribute is modified only when BGP packets traverse point-to-point links.

Q6) Which three statements are true regarding the BGP AS-path attribute? (Choose three.)

A) The local AS number is prepended to the AS path each time that the route crosses an AS boundary.

B) The AS that originally injected the route into BGP is always found in the rightmost end of the AS path.

C) The AS-path attribute is also used to avoid routing loops.

D) BGP routes with an empty AS path were injected into BGP from outside the local AS.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Influencing BGP Route Selection with Weights

Overview

This lesson discuses how to influence Border Gateway Protocol (BGP) route selection by setting the weight attribute of incoming BGP routes. Three methods used to set the weight attribute are discussed in this lesson as follows: default weight, filter-list weight, and setting the weight attribute with route-maps. This lesson also explains how to monitor the BGP table to verify correct weight configuration and properly influenced path selection.

Importance

When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. The “optimum” or “best” route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides many tools for administrators to influence route selection. One of these tools is the weight attribute.

Objectives

Upon completing this lesson, you will be able to:

- List BGP route selection criteria
- Describe the use of BGP weights to influence the BGP route selection process
- Influence the BGP route selection process by configuring per-neighbor weights
- Influence the BGP route selection process by configuring BGP weights with AS-path filters
- Influence the BGP route selection process by configuring BGP weights with route-maps

- Identify the Cisco IOS® commands required to monitor BGP route selection and weights
- Summarize BGP route selection and filtering tools

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

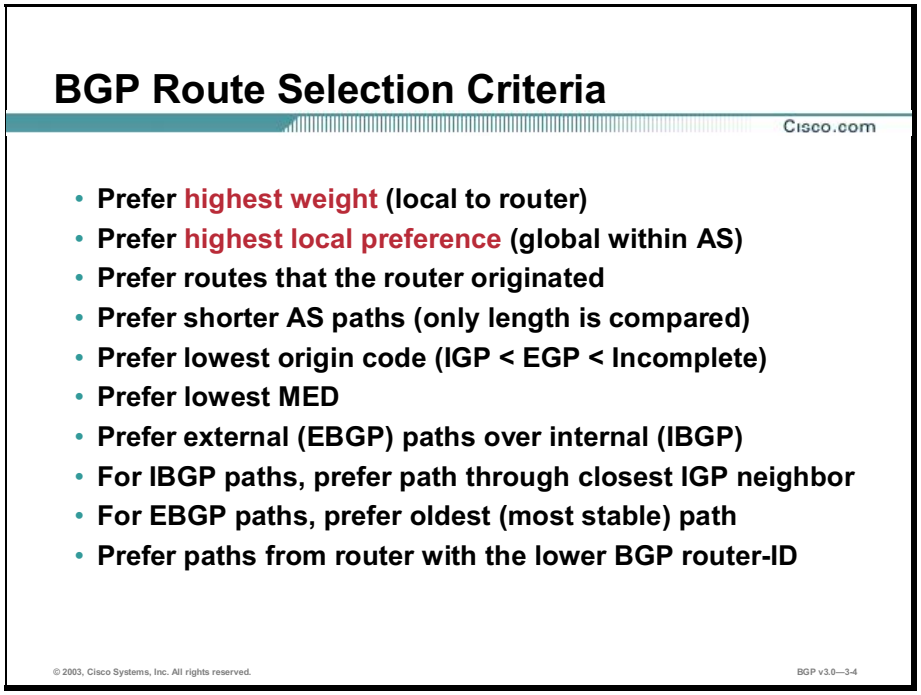
This lesson includes these topics:

- Overview
- BGP Route Selection Criteria
- Influencing BGP Route Selection
- Configuring Per-Neighbor Weights
- Changing Weights with AS-Path Filters
- Changing Weights with Route-Maps
- Monitoring BGP Route Selection and Weights
- BGP Route Selection and Filtering Tools Summary
- Summary
- Assessment (Lab): Influencing BGP Route Selection with Weights

 Copyright © 2003, Cisco Systems, Inc.

BGP Route Selection Criteria

This topic lists the different criteria used by BGP for best-path route selection.


BGP Route Selection Criteria
Cisco.com
• Prefer highest weight (local to router)
• Prefer highest local preference (global within AS)
• Prefer routes that the router originated
• Prefer shorter AS paths (only length is compared)
• Prefer lowest origin code (IGP < EGP < Incomplete)
• Prefer lowest MED
• Prefer external (EBGP) paths over internal (IBGP)
• For IBGP paths, prefer path through closest IGP neighbor
• For EBGP paths, prefer oldest (most stable) path
• Prefer paths from router with the lower BGP router-ID
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-4

BGP route selection criteria take the weight parameter into consideration first. If a router has two alternative paths to the same destination, and their weight values are different, BGP selects the route with the highest weight value as the best. Only when the two alternatives have equal weight is the next criterion, local preference, checked.

A high local preference value is preferred before a low value. Only when the two alternatives have an equal local preference is the next criterion checked.

Practice

Q1) What is the difference between local preference and weight?

A) Local preference has a higher priority in BGP path selection.

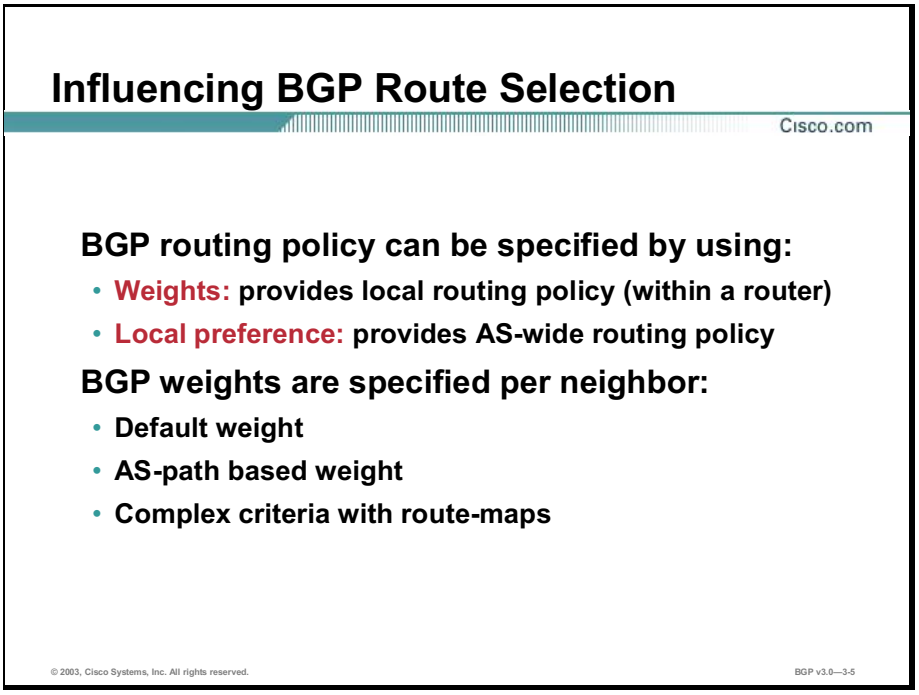
B) Local preference is used AS-wide while weight is local to a single router.

C) Local preference is local only to a specific BGP-speaking router.

D) Local preference is used to influence incoming path selection.

Influencing BGP Route Selection

This topic describes how network administrators can use BGP weights to influence the BGP route selection process.


Influencing BGP Route Selection
Cisco.com
BGP routing policy can be specified by using:
• Weights: provides local routing policy (within a router)
• Local preference: provides AS-wide routing policy
BGP weights are specified per neighbor:
• Default weight
• AS-path based weight
• Complex criteria with route-maps
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-5

The **Weight** attribute is local to a single router only. The weight value is never propagated by the BGP protocol. It constitutes a routing policy local to the router.

Local preference is assigned to a route as an attribute. It is carried with the route on all internal BGP sessions. This situation means that all other BGP-speaking routers within the autonomous system (AS) receive the same information. Normally, a router assigns a local preference to a route when it is received on an external BGP session, before it is accepted and entered in the BGP table of the border router. Routers propagate the local preference attribute on internal BGP sessions only. This policy constitutes a routing policy for the entire AS.

The router can assign the weight attribute to a route three ways:

- All routes received from a specific neighbor can be assigned a weight value. This weight value indicates that the neighbor is more preferred than the other neighbors.
- Received routes from a neighbor can be matched against an AS-path filter. Those matched by the filter are assigned a weight value. Those not matched are accepted, but their weight value is not set.
- A route-map applied on incoming routes from a neighbor can be used to select some routes and assign them weight values. Remember that a route-map also acts as a filter and will silently drop those routes not permitted by any statement in the route-map.

 Copyright © 2003, Cisco Systems, Inc.

If configured, the default weight assignment on routes received from a neighbor is applied first. All routes received from the neighbor are assigned a weight value as defined by the default weight.

Secondly, a filter-list, which does not filter routes out, is applied if configured on the router to select those routes received from the neighbor that should be assigned a specific weight value. Routes permitted by the filter-list have their weight values changed to the value indicated.

Lastly, a route-map is applied, if configured on the router. The route-map can be arbitrarily complex and select routes based on various selection criteria, such as a network number or AS path. The selected routes can have some attributes altered. The route-map can set the weight values of permitted routes. Selection can be done in several route-map statements, giving the opportunity to assign a certain weight value to some routes and another weight value to others. A route-map can also completely filter out routes.

Practice

Q1) What three methods can you use to set the BGP weight attribute? (Choose three.)

A) route-map

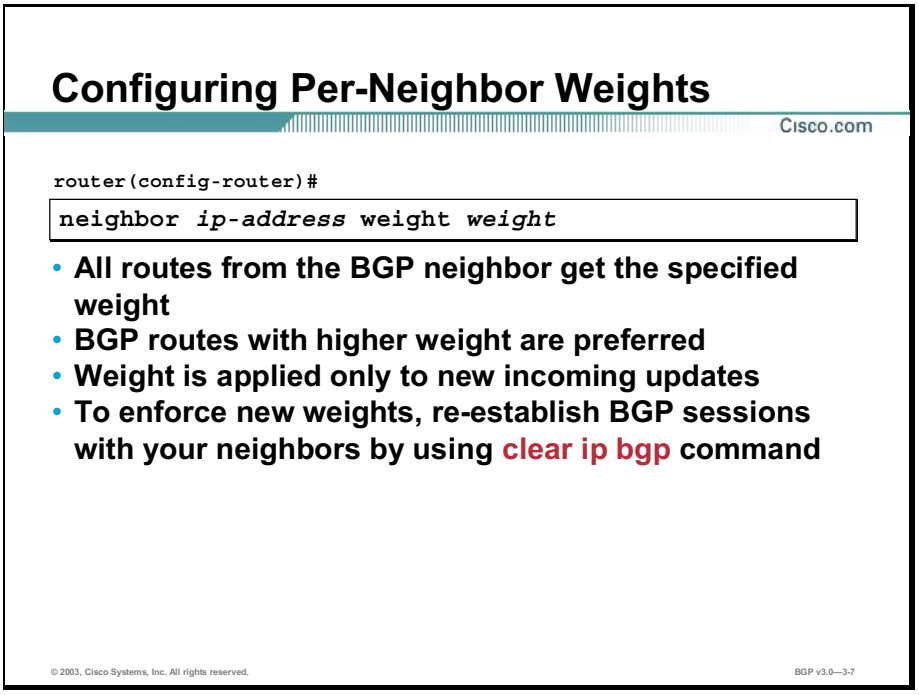
B) AS-path filter-list

C) access control list

D) default weight assigned to a specific neighbor

Configuring Per-Neighbor Weights

This topic describes how to influence the BGP route selection process by configuring per-neighbor weights.


Configuring Per-Neighbor Weights
Cisco.com
router(config-router)#
neighbor ip-address weight weight
• All routes from the BGP neighbor get the specified weight
• BGP routes with higher weight are preferred
• Weight is applied only to new incoming updates
• To enforce new weights, re-establish BGP sessions with your neighbors by using clear ip bgp command
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-7

neighbor weight

To assign a weight to a neighbor connection, use the **neighbor weight** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **weight** *weight*

To remove a weight assignment, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **weight** *weight*

Syntax Description

ip-address	IP address of neighbor.
peer-group-name	Name of a BGP peer group.
weight	Weight to assign. Acceptable values are 0 to 65535.

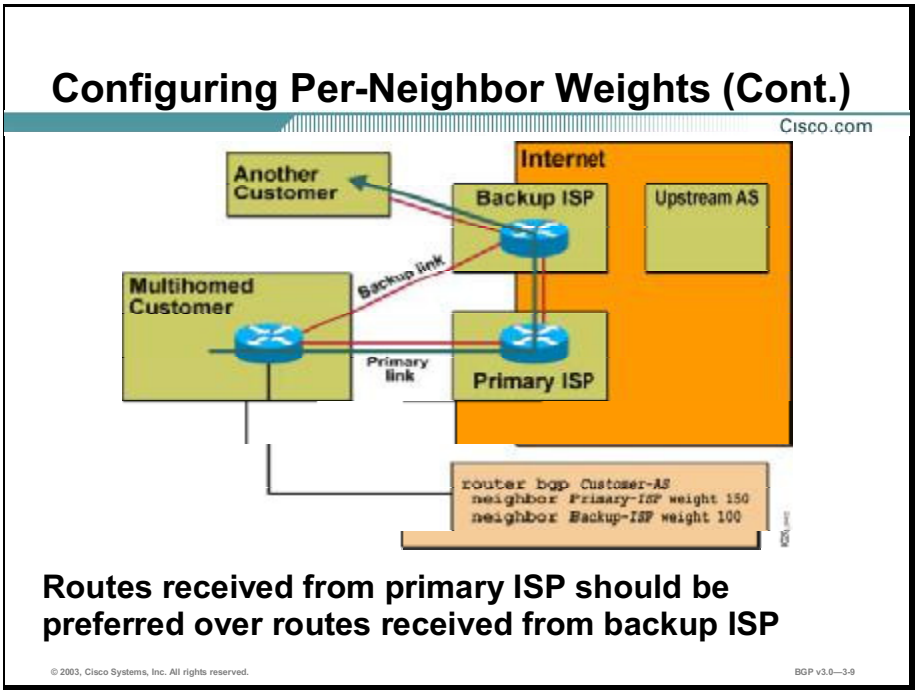
 Copyright © 2003, Cisco Systems, Inc.

All routes received from the neighbor after the configuration line is in place are assigned the weight value. To make sure that all routes from the neighbor receive the new weight value, you can restart the BGP session, thus forcing the neighbor to resend all routes.

If no weight value is specified, the default value of 0 is applied.

Restarting of BGP sessions might be necessary after making a configuration change in the routing policy. The configuration change itself will not alter the already-received routes. The **clear ip bgp** EXEC command tears down the BGP session, and the session automatically restarts.

Example

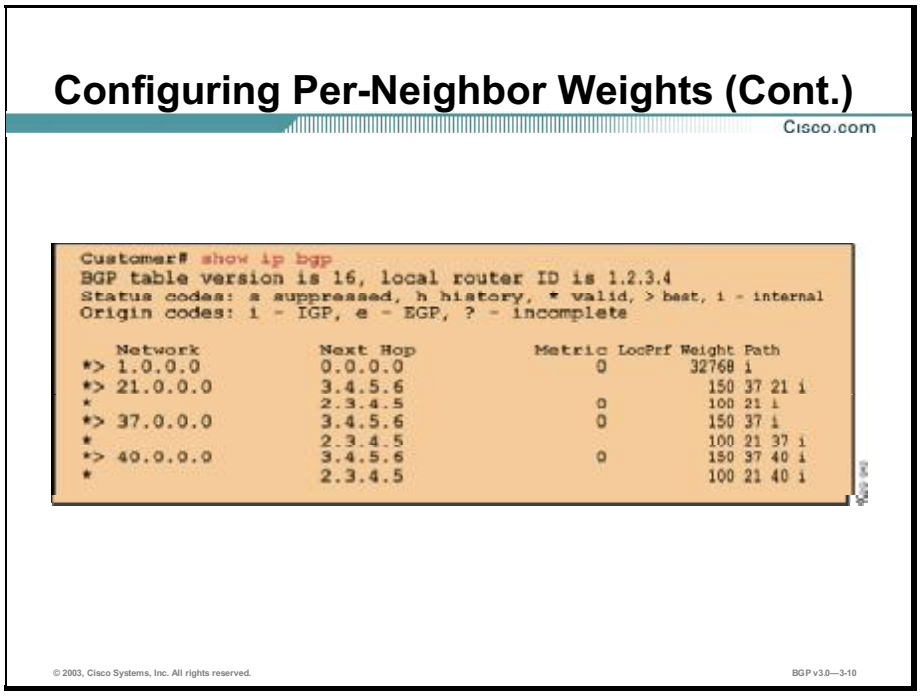


In this example, the multihomed customer would like to use the primary link to the primary Internet service provider (ISP) for all destinations. The weight is configured by the customer on both BGP sessions, giving a higher weight to those routes received from the primary ISP compared to those received from the backup ISP.

Any time the multihomed customer receives routing information about the same IP network number from both the ISPs, the customer compares the weights assigned to the routes. Those received from the primary ISP will always win this comparison. The multihomed customer sends the outgoing IP packets to the destination network via the primary ISP regardless of the other BGP attributes assigned to both alternatives.

Consequently, the other customer directly connected to the backup ISP will also be reached via the primary ISP.

 Copyright © 2003, Cisco Systems, Inc.


Configuring Per-Neighbor Weights (Cont.)
Cisco.com
Customer# show ip bgp
BGP table version is 16, local router ID is 1.2.3.4
Status codes: s suppressed, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 1.0.0.0 0.0.0.0 0 32768 i
*> 21.0.0.0 3.4.5.6 150 37 21 i
* 2.3.4.5 0 100 21 i
*> 37.0.0.0 3.4.5.6 0 150 37 i
* 2.3.4.5 100 21 37 i
*> 40.0.0.0 3.4.5.6 0 150 37 40 i
* 2.3.4.5 100 21 40 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-10

In this example, the multihomed customer has received routes to three different class A networks outside of its own AS (network 21.0.0.0/8, network 37.0.0.0/8, and network 40.0.0.0/8). The customer received all three routes from both the primary ISP and the backup ISP.

When the routes were received from the primary ISP, the weight value 150 was assigned to each of the routes. When the routes were received from the backup ISP, the weight value 100 was assigned to each of the routes.

The customer router now makes the route selection. It has two alternative paths for each destination network. For each of them, the router selects the path via the primary ISP as the best. It makes this selection regardless of other BGP attributes, such as AS-path length.

The network 21.0.0.0/8 is reached via the primary ISP although it is actually a network in the AS of the backup ISP (AS 21).

The class A network 1.0.0.0/8 in this example is injected into the BGP table by this router. By default, locally sourced routes are assigned a weight of 32768.

Practice

Q1) What is the default weight for routes received from a BGP neighbor?

A) 0

B) 100

C) 32768

D) depends on the Cisco IOS release

Q2) What default weight is applied to locally sourced BGP routes?

A) 0

B) 100

C) 32768

D) depends on the Cisco IOS release

Q3) When are the weights configured on a neighbor enforced?

A) Before the new weights can take effect, the BGP process on the router must be removed and reconfigured.

B) The router must first be rebooted for the new weights to take effect.

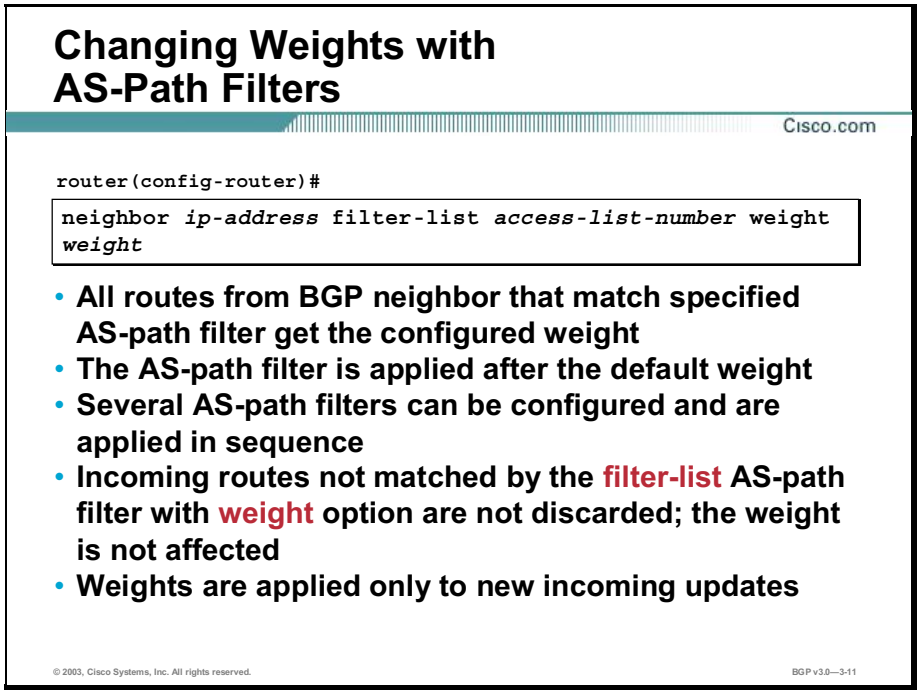
C) The new weights will be applied after the BGP update interval of 30 minutes expires.

D) The new weight configuration is applied to all routes received following the configuration change.

 Copyright © 2003, Cisco Systems, Inc.

Changing Weights with AS-Path Filters

This topic describes how to influence the BGP route selection process by configuring BGP weights with AS-path filters.


Changing Weights with AS-Path Filters
Cisco.com
router(config-router)#
neighbor ip-address filter-list access-list-number weight weight
• All routes from BGP neighbor that match specified AS-path filter get the configured weight
• The AS-path filter is applied after the default weight
• Several AS-path filters can be configured and are applied in sequence
• Incoming routes not matched by the filter-list AS-path filter with weight option are not discarded; the weight is not affected
• Weights are applied only to new incoming updates
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-11

neighbor filter-list

To set up a BGP filter, use the **neighbor filter-list** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **filter-list** *access-list-number* {**in** | **out** | **weight** *weight*}

To disable this function, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **filter-list** *access-list-number* {**in** | **out** | **weight** *weight*}

Syntax Description

ip-address	IP address of the neighbor.
peer-group-name	Name of a BGP peer group.
access-list-number	Number of an AS-path access list. You define this access list with the **ip as-path access-list** command.
in	Access list to incoming routes.
out	Access list to outgoing routes.

weight *weight* Assigns a relative importance to incoming routes matching AS paths. Acceptable values are 0 to 65535.

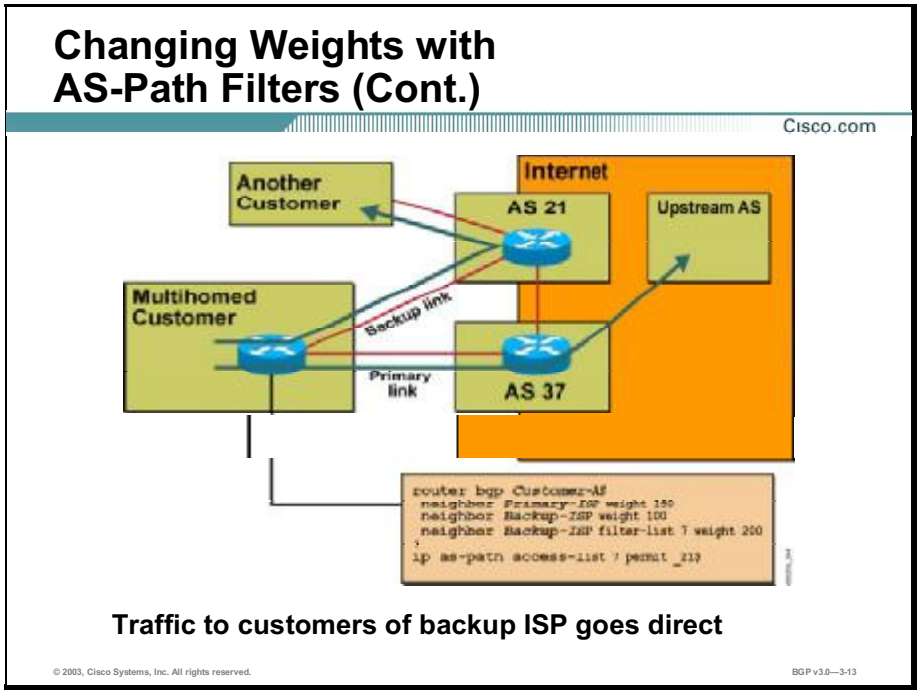
When you use the keyword **weight**, the router does not use the filter-list to discard any routes. Instead, the filter-list is applied on the incoming routes and selects those that should have their weight value altered. Those that the filter-list does not select are accepted without changing the weight.

Restarting of BGP sessions might be necessary after configuring a change in the routing policy. BGP applies the new configuration only on routes coming in after the configuration change.

Note Specifying weights with filter-lists is no longer supported in Cisco IOS Release 12.1, and the command has already been removed from Cisco IOS release 12.1T. These releases use an incoming route-map, where you match an AS path with the **match as-path** command and set weight with the **set weight** command. When using a route-map as a replacement for the filter-list with **weight** option, make sure that specifying a “permit” entry in the route-map without an associated match condition does not filter all other routes. Using route-maps as a weight-setting mechanism is explained later in this lesson.

 Copyright © 2003, Cisco Systems, Inc.

Example

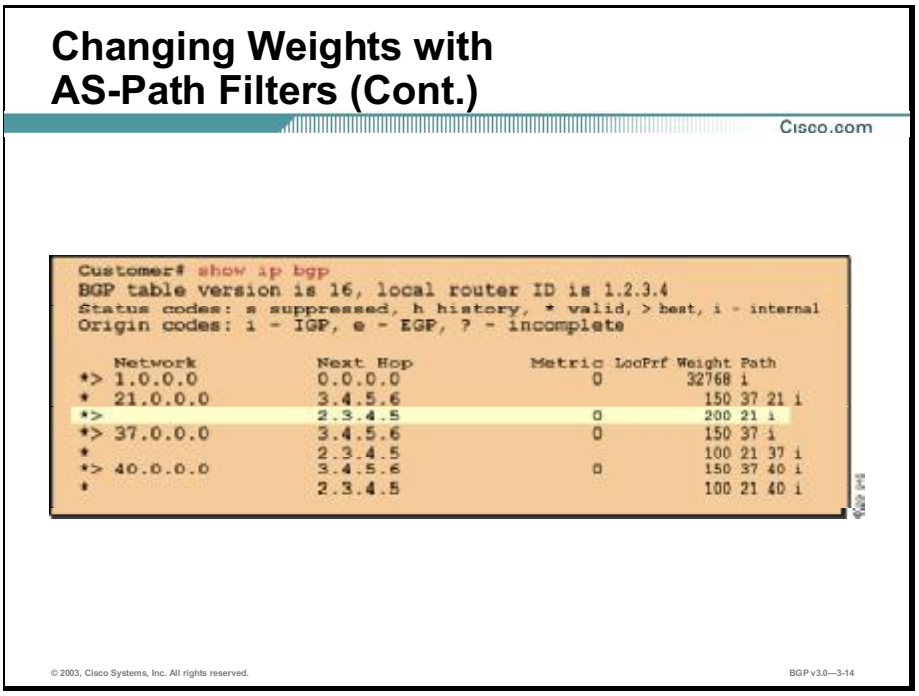


Sometimes a blind preference to use a specific provider in a multihomed network can cause strange results. The case where a customer reaches networks in the AS of the backup ISP via the primary ISP even though it has a direct connection to the backup ISP is an example of a network configuration leading to strange results.

To change BGP path selection into a more predictable behavior, the blind preference has to be replaced with a more selective preference. In the example, the backup ISP has the AS number 21. Within that AS there are many single-homed customer networks that do not have their own AS number but are part of the AS of the backup ISP.

All routes received from the primary ISP are assigned the weight 150 by the customer edge router. But routes received from the backup ISP are assigned weight values in a more selective way. First, all routes are assigned the value 100. Then those routes selected by the AS-path access-list 7 have their values changed. The filter-list 7 selects those routes having an AS path containing only 21, indicating that they were created in AS 21.

The result is that most routes are preferred to reach the network via the primary ISP. Only those routes created in the AS of the backup ISP are preferred to reach the network via the direct link to the backup ISP.


Changing Weights with
AS-Path Filters (Cont.)
Cisco.com
Customer# show ip bgp
BGP table version is 16, local router ID is 1.2.3.4
Status codes: s suppressed, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 1.0.0.0 0.0.0.0 0 32768 i
* 21.0.0.0 3.4.5.6 150 37 21 i
*> 2.3.4.5 0 200 21 i
*> 37.0.0.0 3.4.5.6 0 150 37 i
* 2.3.4.5 100 21 37 i
*> 40.0.0.0 3.4.5.6 0 150 37 40 i
* 2.3.4.5 100 21 40 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-14

The **show ip bgp** command output shows the BGP table of a multihomed customer in AS 123. It has direct links to two different ISPs. The primary ISP is AS 37, and the backup ISP is AS 21.

In this example, the customer would like to reach the destination networks within AS 21 directly via the backup ISP. To accomplish this, the customer must configure a more selective assignment of weights. First, the customer assigns the default weight 150 to all routes received from the primary ISP and the weight 100 to all routes received from the backup ISP. Additionally, the routes received from the backup ISP must be checked by a filter-list. Those permitted by the list have their weight values changed to 200. The filter-list selects those routes having an AS path indicating that they were created in AS 21.

As a result, the customer reaches all networks outside its own AS via the primary ISP, except for the class A network 21.0.0.0/8, which it reaches via the backup ISP.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) How could you implement a primary/backup ISP routing policy using weights?

A) Assign higher weights to all routes received from the backup ISP.

B) Assign lower weights to all routes received from the backup ISP.

C) Assign higher weights to all routes received from the primary ISP.

D) Assign lower weights to all routes received from the primary ISP.

Q2) What is the difference between a **filter-list in** configuration command and a **filter-list weight** configuration command?

A) The filter-list without the **weight** keyword can manipulate AS paths or the weight attribute, but the **weight** keyword is limited to the weight attribute.

B) **Filter-list weight** will not filter out any route but will assign the weight value to those routes permitted by the filter-list.

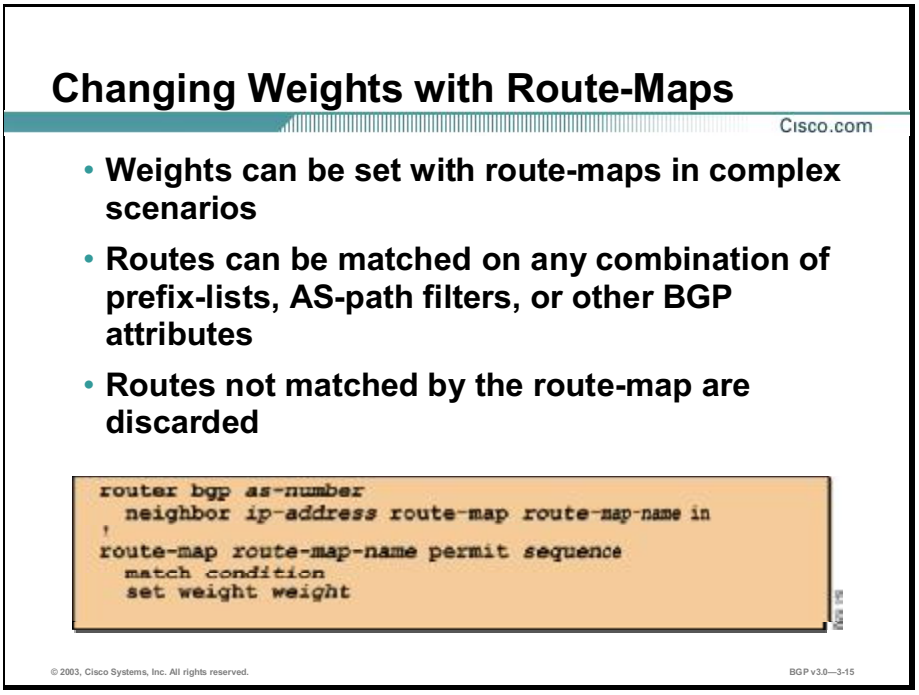
C) Using **filter-list weight** forces the router to use process switching.

D) **Filter-list weight** can match an AS path with a single AS number entry only.

Copyright © 2003, Cisco Systems, Inc.

Changing Weights with Route-Maps

This topic describes how to influence the BGP route selection process by configuring BGP weights with route-maps.


Changing Weights with Route-Maps
Cisco.com
• Weights can be set with route-maps in complex scenarios
• Routes can be matched on any combination of prefix-lists, AS-path filters, or other BGP attributes
• Routes not matched by the route-map are discarded
router bgp as-number
neighbor ip-address route-map route-map-name in
!
route-map route-map-name permit sequence
match condition
set weight weight
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-15

The route-map is a powerful tool to select and alter routing information. When a route-map is applied to incoming information from a BGP neighbor, each received update is examined as it passes through the route-map. Statements in the route-map are executed in the order specified by their sequence numbers

The first statement in the route-map that has all the match clauses indicating a match is the one used. If the route-map says “permit,” the set clauses are applied to the route, the route is accepted, and the weight is changed.

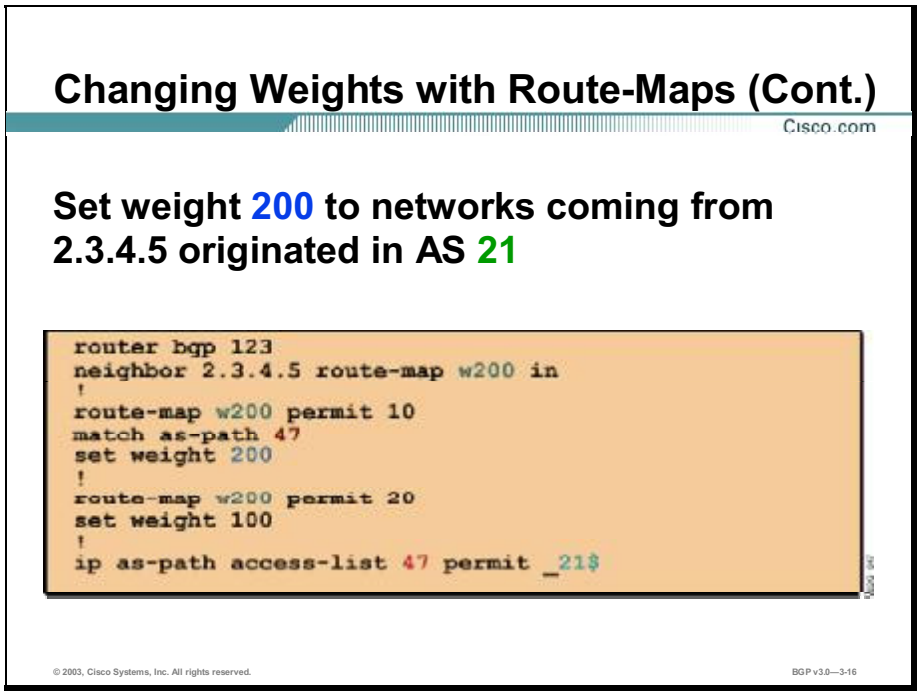
Match clauses can be arbitrarily complex. One of them can refer to an AS-path access-list that does matching on AS paths. Another can refer to a prefix-list that does matching on the announced network number. Only when all configured match clauses permit is the route-map statement used and its result, “permit” or “deny,” applied.

If a received route is not matched by any of the route-map statements, and the end of the route-map is reached, the route-map logic has an “implicit deny” rule. This rule means that if no statement selects a route, the route is discarded.

If the “implicit deny” rule is not desired, an “explicit permit all” at the end of the route-map can overrule it. To ensure that such a route-map statement is the last statement, you should assign it a very high sequence number. It should not have any match clause at all. The lack of a match clause means, “match all”. By not configuring any set clause, you can ensure that no attributes are altered by the statement.

 Copyright © 2003, Cisco Systems, Inc.

Example



This is an example of a route-map that sets the weight value to each route received from a neighbor.

All received routes have their AS paths checked against the AS-path access-list 47. Those routes having an AS path indicating that originated in AS 21 are permitted by the AS-path access-list 47 as referenced by route-map statement number 10. Routes permitted and selected by route-map statement number 10 in the w200 route-map will have their weight set to 200 as indicated by the set clause in the route-map.

The routes that are not originated in AS 21 (routes not permitted by AS-path access-list 47) are then tested by route-map statement number 20. This statement does not include a match clause, indicating that all routes are matched. Therefore, all routes not matched by route-map statement 10 are matched by route-map statement 20. The route-map has been configured with an "explicit permit all" statement at the end of the route-map.

Routes matched by route-map statement 20 have their weight set to 100. The result is that the routes originated in AS 21 are accepted by the router and assigned the weight 200. All others are accepted and assigned the weight value 100. No route is discarded by this route-map.

Practice

Q1) When you are using route-maps to modify weights, what happens by default to a route that does not match any of the route-map statements?

A) The route is accepted with the weight attribute unmodified.

B) The route is discarded.

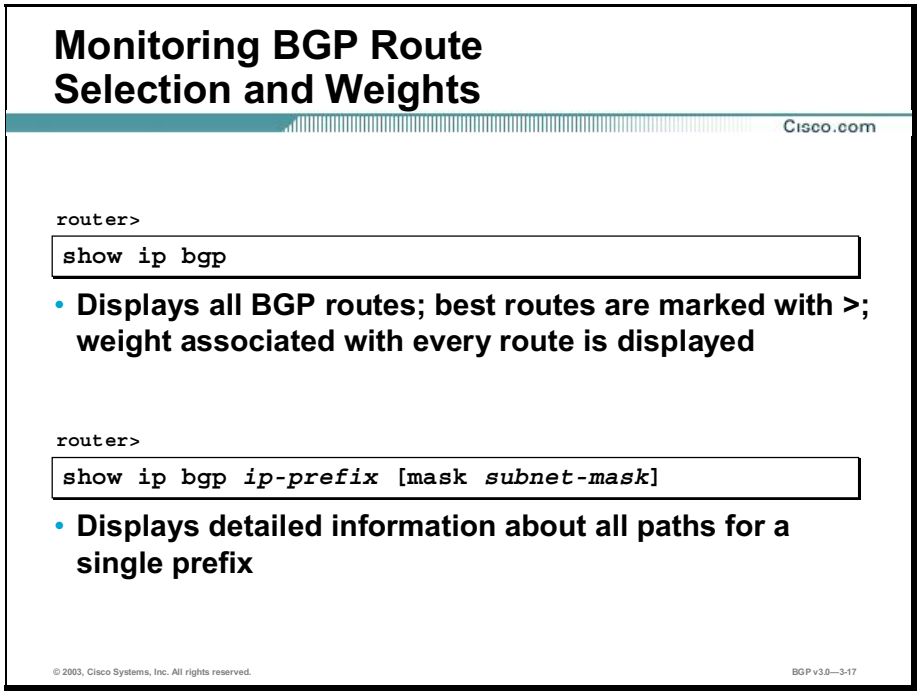
C) The route will be inserted into the BGP table but not the IP routing table.

D) An error will be displayed on the router console and in router debugs.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring BGP Route Selection and Weights

This topic lists the Cisco IOS commands required to monitor BGP route selection and weights.


Monitoring BGP Route Selection and Weights
Cisco.com
router>
show ip bgp
• Displays all BGP routes; best routes are marked with >; weight associated with every route is displayed
router>
show ip bgp ip-prefix [mask subnet-mask]
• Displays detailed information about all paths for a single prefix
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-17

show ip bgp

To display entries in the BGP routing table, use the **show ip bgp** EXEC command.

show ip bgp [*network*] [*network-mask*] [**longer-prefixes**]

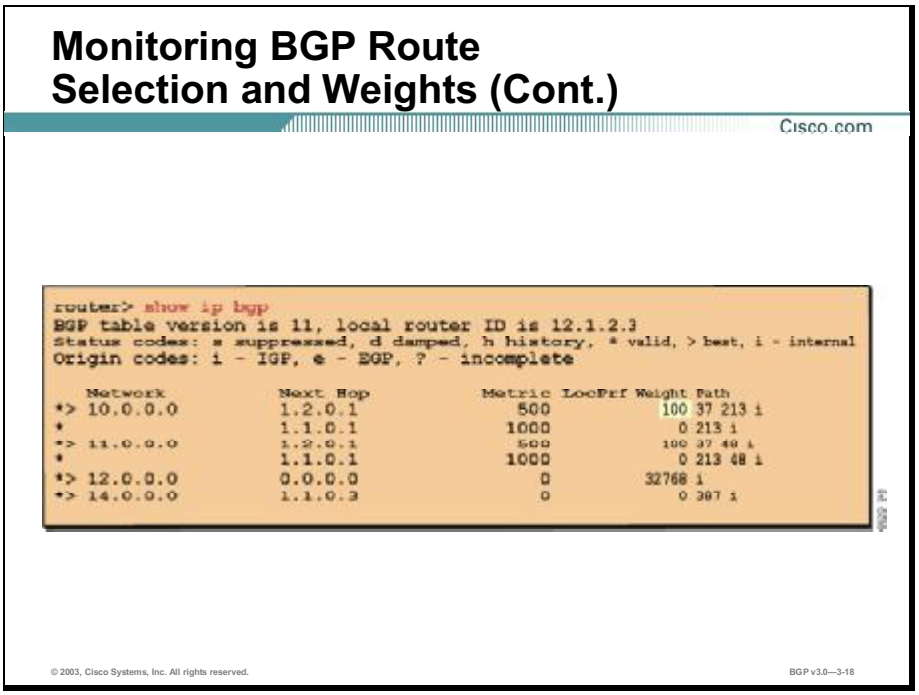
Syntax Description

network	(Optional) Network number, entered to display a particular network in the BGP routing table
network-mask	(Optional) Displays all BGP routes matching the address/mask pair
longer-prefixes	(Optional) Displays route and more specific routes

Without any argument, the **show ip bgp** command displays the entire BGP table. The routes selected as the best are indicated by the “>” character.

To get more detailed information about routes to a specific destination network, you can use the network number, and optionally the subnet mask, as an argument on the command line. These additions will display more detailed information about that specific network.

Example


Monitoring BGP Route Selection and Weights (Cont.)
Cisco.com
router> show ip bgp
BGP table version is 11, local router ID is 12.1.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 10.0.0.0 1.2.0.1 500 100 37 213 i
* 1.1.0.1 1000 0 213 i
*> 11.0.0.0 1.2.0.1 500 100 37 48 i
* 1.1.0.1 1000 0 213 48 i
*> 12.0.0.0 0.0.0.0 0 32768 i
*> 14.0.0.0 1.1.0.3 0 0 387 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-18

The **show ip bgp** command gives a printout of all routes in the BGP table. Each route is displayed on one line. This one-line limitation means that more detailed information about the route cannot be displayed, due to lack of space.

The network number is displayed, and if the subnet mask differs from the natural mask, the prefix length is indicated. The BGP next-hop attribute, multi-exit discriminator (MED; metric), local preference, weight, AS-path, and origin code are displayed on the line. Local preference is displayed only if it is not the default value.

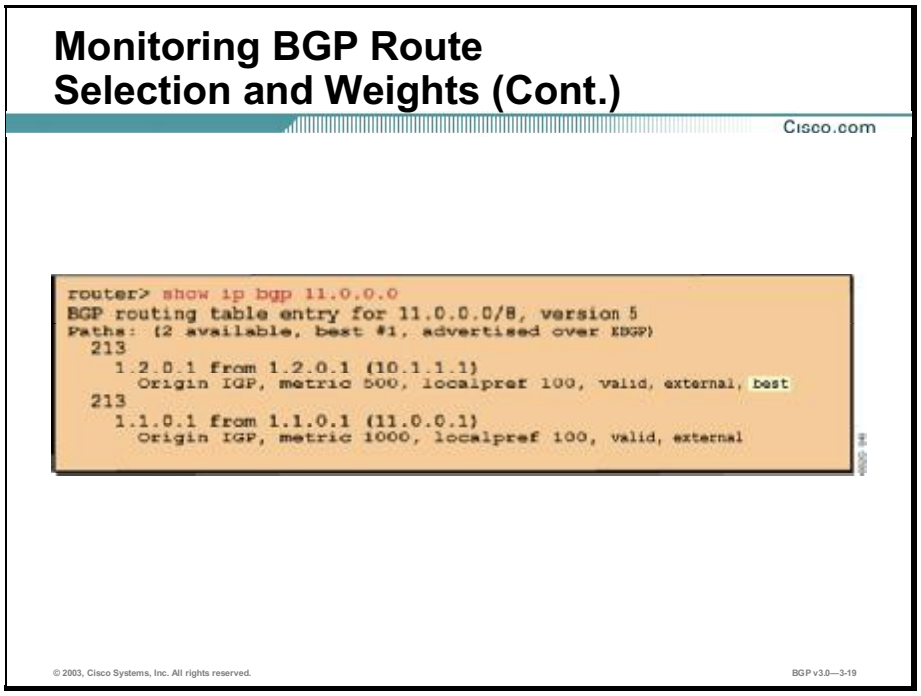
The printout is sorted in network number order. If there is more than one route to the same network, the network number is printed on the first line only. The other routes to the same network have their network field left blank on the output.

Routes selected as the best to reach a certain destination network are indicated by the “>” character.

In this example, weight has been used to prefer routes received from the neighbor in AS 37. Therefore, although the AS path is shorter via AS 213, the class A network 10.0.0.0/8 is reached via AS 37 (because the weight is higher).

Information about network 14.0.0.0/8 is received only from the neighbor in AS 387. As there is no alternative, the route is selected as best.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring BGP Route Selection and Weights (Cont.)
Cisco.com
router> show ip bgp 11.0.0.0
BGP routing table entry for 11.0.0.0/8, version 5
Paths: (2 available, best #1, advertised over EBGP)
213
1.2.0.1 from 1.2.0.1 (10.1.1.1)
Origin IGP, metric 500, localpref 100, valid, external, best
213
1.1.0.1 from 1.1.0.1 (11.0.0.1)
Origin IGP, metric 1000, localpref 100, valid, external
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-19

The **show ip bgp** command with network number as argument displays more detailed information about that network only. First, a short summary indicating the network number and prefix length is displayed along with the table version number for this route. The next line says how many alternative routes have been received, and which one of them has been selected by the router as the best.

Next, there are a couple of lines for each of the received routes to reach the network. For each of the routes, all attributes are displayed. The one selected as the best also has the word “best” displayed.

In this example, there are two alternatives to reach network 11.0.0.0/8. Each of them is received from different neighbors in AS 213. The network 11.0.0.0 is created in AS 213.

The route selection mechanism has selected the first route listed as the best. It was chosen because the MED (metric) value is lower.

Practice

Q1) What three pieces of information can you obtain from the output of the **show ip bgp** command? (Choose three.)

A) the best route to a destination

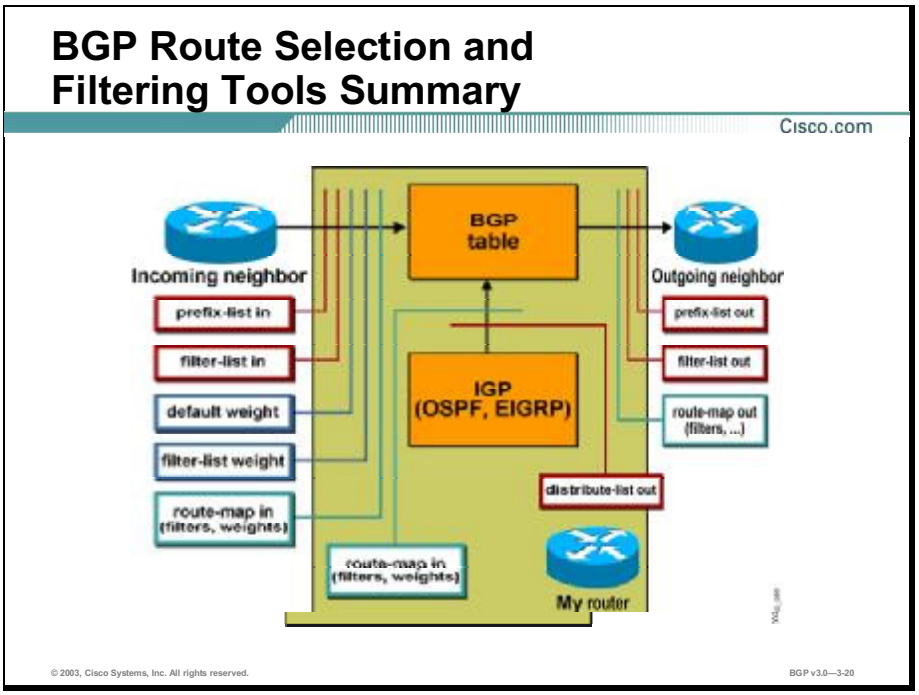
B) the weight attribute

C) the AS path of the route

D) the administrative distance of the route

BGP Route Selection and Filtering Tools Summary

This topic presents a summary of all BGP filtering tools in the order they are applied.



The figure shows all the possible applications of prefix-lists, filter-lists, weights, and route-maps. They are applied in the order indicated.

Prefix-lists and filter-lists, both in and out, filter out routes and discard those not permitted. Weight setting is applicable only on incoming routes because a router never propagates the weight attribute to its neighbors. Route-maps can be filters that discard routes but can also be used to modify and set various attributes on both incoming and outgoing routes.

Practice

Q1) Which method of influencing route selection with weights is the last to be applied on an incoming interface?

A) prefix-list

B) route-map

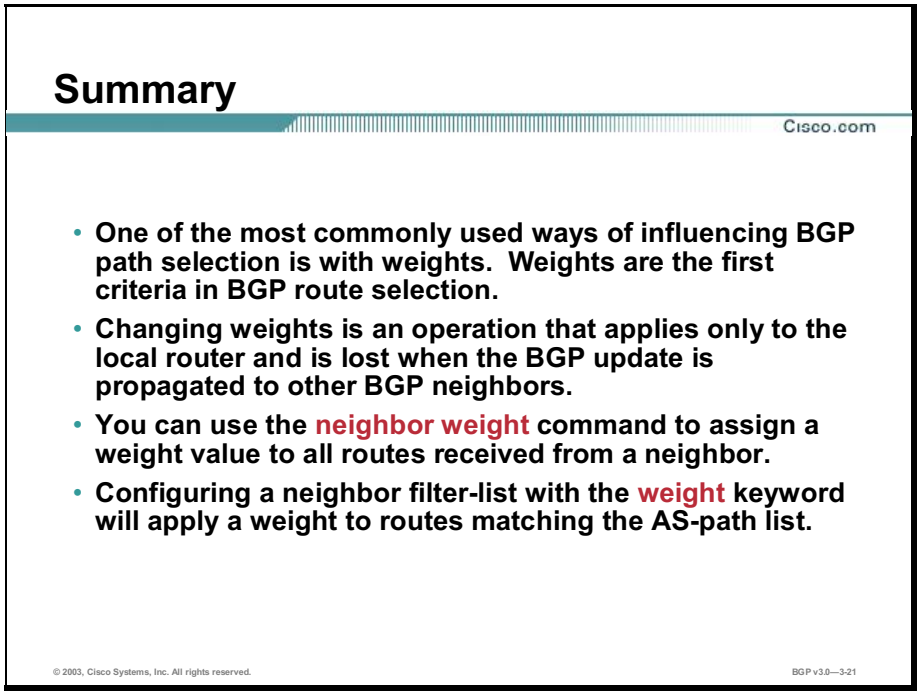
C) filter-list weight

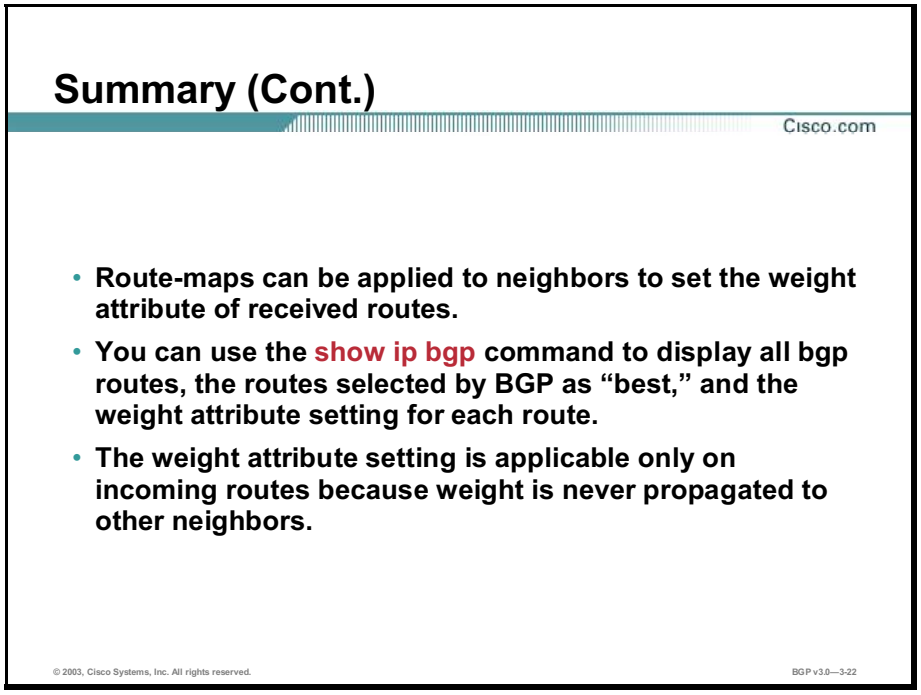
D) default weight

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• One of the most commonly used ways of influencing BGP path selection is with weights. Weights are the first criteria in BGP route selection.
• Changing weights is an operation that applies only to the local router and is lost when the BGP update is propagated to other BGP neighbors.
• You can use the neighbor weight command to assign a weight value to all routes received from a neighbor.
• Configuring a neighbor filter-list with the weight keyword will apply a weight to routes matching the AS-path list.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-21


Summary (Cont.)
Cisco.com
• Route-maps can be applied to neighbors to set the weight attribute of received routes.
• You can use the show ip bgp command to display all bgp routes, the routes selected by BGP as “best,” and the weight attribute setting for each route.
• The weight attribute setting is applicable only on incoming routes because weight is never propagated to other neighbors.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-22

Next Steps

After completing this lesson, go to:

- BGP Local Preference lesson

References

For additional information, refer to these resources:

- For more information on the weight attribute, refer to “Border Gateway Protocol” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm
- For further information on the weight attribute, refer to “Configuring BGP” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2/1cfbgp.htm#xtocid15

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: Influencing BGP Route Selection with Weights

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using the weight attribute in a situation where you must support connections to multiple ISPs.

After completing this exercise, you will be able to:

- Influence the BGP route selection process by configuring per-neighbor weights
- Influence the BGP route selection process by configuring BGP weights with route-maps
- Monitor BGP route selection and weights

Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode
neighbor *ip-address* **weight** *weight*	Assign weight to all received updates from specified neighbor
neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in | out**}	To apply a route-map to incoming or outgoing routes
route-map *name* {**permit | deny**} **seq**	Define or modify an existing entry in a route-map
match as-path *list-number*	Specify route-map matching criteria to match an as-path access list
set weight *weight*	Set weight in a route-map
show ip bgp summary	Verify if both BGP sessions are up
show ip bgp	Inspect the contents of the BGP table
clear ip bgp	Clear the BGP session with your neighbor

Job Aids

These job aids are available to help you complete the laboratory exercise:

- Currently, you are using the service provider “Cheap” as your primary provider for Internet connectivity and the service provider “Good” as your backup provider. As the result of this policy, the link toward the “Good” service provider is underused, while the link toward the provider “Cheap” is overloaded.

- An analysis of the Internet structure beyond your service providers indicates that you could improve the link use if you send traffic toward AS 213 and AS 37 directly to the “Good” service provider.

- In this exercise, you will improve the simple routing policy created in the Multihomed BGP Networks lab exercise, by specifying BGP weights with route-maps.

- You must implement the following routing policy:

  - Prefer routes announced from router “Cheap” over those announced from router “Good”

  - Prefer routes going through or originating in AS 213 or AS 37 from router “Good”

- Figure 1 shows the connectivity that is established between your AS and the two service providers “Good” and “Cheap.”

 Copyright © 2003, Cisco Systems, Inc.

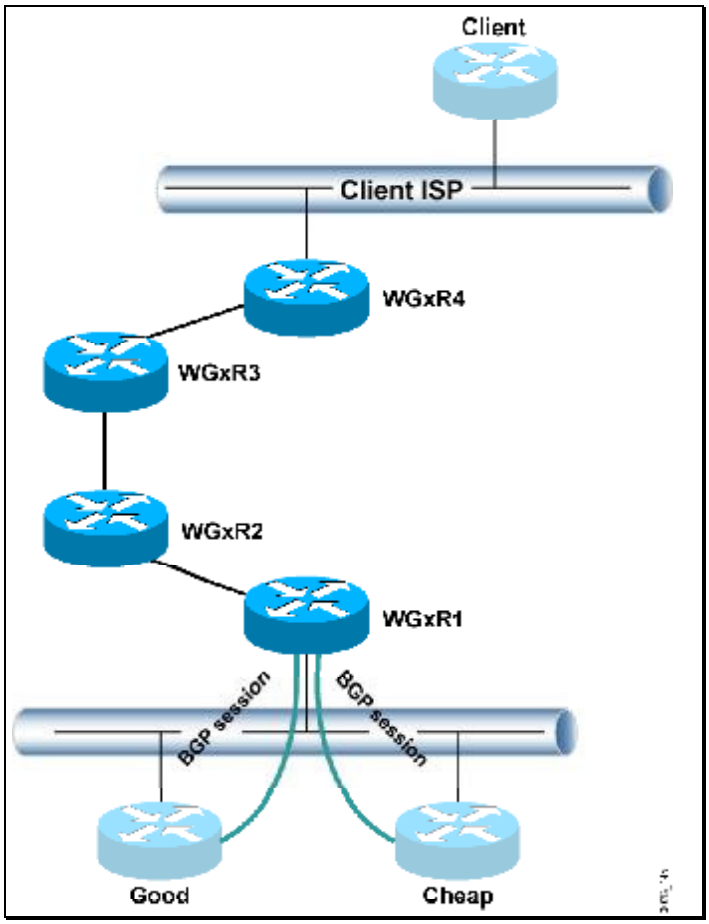


Figure 1: Connections to two different service providers

Exercise Procedure

Complete these steps:

Configuring AS-path access-lists:

Step 1 Create an AS-path access-list that permits AS numbers 213 and 37 in the AS path.

Setting per-neighbor default weights:

Step 2 Set default weights for both neighboring routers.

Step 3 Inspect your BGP table to verify that you prefer routes coming from “Cheap” to routes coming from “Good.”

Using AS-path filters within route-maps to set weights:

Step 4 Create a new route-map. Use the previously configured AS-path access-list as the **match** condition in one of the route-map statements, and set the weight of matched routes as needed.

Step 5 Apply the route-map to incoming updates from router “Good.”

Copyright © 2003, Cisco Systems, Inc.

Exercise Verification

You have completed this exercise when you attain these results:

- Verify your BGP table to see if all prefixes with two paths prefer the one through router "Cheap."

```
wg1r1#sh ip bgp
BGP table version is 43, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  128.20.0.0       192.168.20.20            0           100 20 i
*>                  192.168.20.22            0           200 22 i
*> 128.20.12.0/24   192.168.20.20            0           100 20 i
*  128.22.0.0       192.168.20.20            0           100 20 i
*>                  192.168.20.22            0           200 22 i
*> 128.22.12.0/24   192.168.20.22                        100 20 22 i
*  128.26.0.0       192.168.20.20            0           100 20 42 26
i
*>                  192.168.20.22            0           200 22 26 i
*> 128.37.0.0       192.168.20.20            0           300 20 42 37
i
*                   192.168.20.22            0           200 22 26 42
37 i
*  128.42.0.0       192.168.20.20            0           100 20 42 i
*>                  192.168.20.22            0           200 22 26 42
i
*  128.51.0.0       192.168.20.20            0           100 20 42 26
51 i
*>                  192.168.20.22            0           200 22 26 51
i
*> 128.213.0.0      192.168.20.20            0           300 20 213 i
*> 128.214.0.0      192.168.20.22            0           200 22 214 i
*> 192.20.11.0      192.168.20.22            0           200 22 i
*> 192.20.12.0/30   192.168.20.20                        200 22 20 i
*> 192.22.11.0      192.168.20.22            0           200 22 i
*> 192.22.12.0/30   192.168.20.22            0           200 22 i
*> 192.26.11.0      192.168.20.22            0           200 22 26 i
*> 192.37.11.0      192.168.20.22            0           200 22 26 42
37 i
*> 192.42.11.0      192.168.20.22            0           200 22 26 42
i
*> 192.51.11.0      192.168.20.22            0           200 22 26 51
i
*> 192.168.1.0      0.0.0.0                  0         32768 i
*> 192.214.11.0     192.168.20.22            0           200 22 214 i
*> 197.1.0.0/21     0.0.0.0                           32768 i
*> 197.1.0.0/16     0.0.0.0                           32768 i
s> 197.1.1.0        0.0.0.0                  0         32768 i
s> 197.1.2.0        0.0.0.0                  0         32768 i
```

 Copyright © 2003, Cisco Systems, Inc.

```
s> 197.1.3.0        0.0.0.0                  0         32768 i
s> 197.1.4.0        0.0.0.0                  0         32768 i
s> 197.1.5.0        0.0.0.0                  0         32768 i
s> 197.1.6.0        0.0.0.0                  0         32768 i
s> 197.1.7.0        0.0.0.0                  0         32768 i
s> 197.1.8.0        0.0.0.0                  0         32768 i
*> 197.1.8.0/22     0.0.0.0                            32768 i
*  200.20.0.0/16    192.168.20.20            0           100 20 i
*>                  192.168.20.20                        200 22 20 i
*  200.22.0.0/16    192.168.20.22                        100 20 22 i
*>                  192.168.20.22            0           200 22 i
```

Answer these questions:

Q1) Did all paths automatically get a weight of 100 or 200? Why not? What did you have to do?

Q2) Name some parameters and attributes used for best-path selection.

 Copyright © 2003, Cisco Systems, Inc.

BGP Local Preference

Overview

This lesson discusses how to influence Border Gateway Protocol (BGP) route selection by setting the BGP local preference attribute of incoming BGP routes. Local preference is similar to the weight attribute but differs from the BGP weight attribute in that weight is local to a specific router on which it is configured. Two methods used to set the local preference attribute are discussed in this lesson as follows: default local preference and setting the local preference attribute with route-maps. This lesson also explains how to monitor the BGP table to verify correct local preference configuration and properly influenced path selection.

Importance

When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. The “optimum” or “best” route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides many tools for administrators to influence route selection. One of these tools is the local preference attribute.

Objectives

Upon completing this lesson, you will be able to:

- Explain why using BGP weights may not provide consistent BGP route selection in an AS
- Describe how the BGP local preference attribute influences BGP route selection
- Identify the Cisco IOS® commands required to configure default BGP local preference on a router

- Identify the Cisco IOS commands required to configure BGP local preference using route-maps
- Identify the Cisco IOS commands required to monitor BGP local preference

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

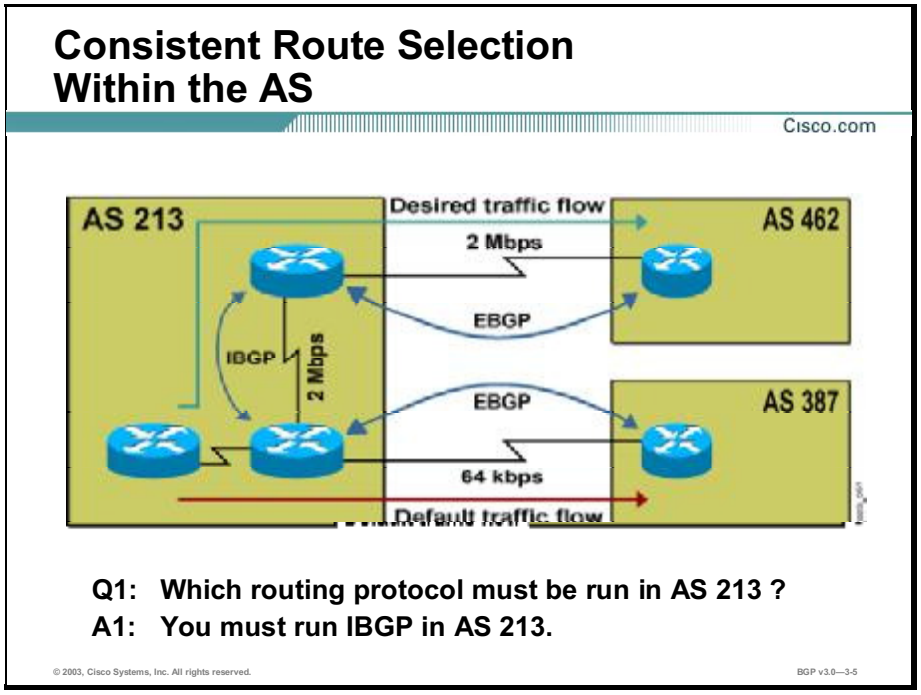
This lesson includes these topics:

- Overview
- Consistent Route Selection Within the AS
- BGP Local Preference
- Configuring Default Local Preference
- Configuring Local Preference with Route-Maps
- Monitoring Local Preference
- Summary
- Assessment (Lab): BGP Local Preference

 Copyright © 2003, Cisco Systems, Inc.

Consistent Route Selection Within the AS

This topic explains why using BGP weights may not provide consistent BGP route selection inside of an autonomous system (AS).

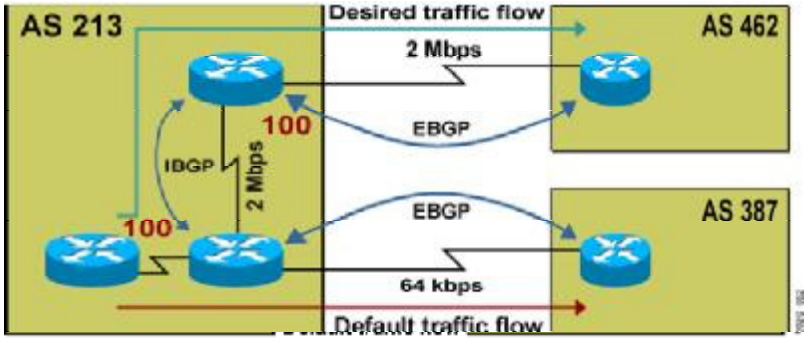


Using BGP in autonomous systems with a single neighbor relationship usually does not require any advanced features. In situations like the one shown in the figure, however, it is important to ensure that customer routers choose the correct link. Obviously, the router should choose the 2-Mbps link and use the 64-kbps link only for backup purposes.

To make sure that the router selects the upper link (2-Mbps link) as its primary link, and has the ability to switch over to the backup if a failure occurs, you must configure an internal BGP session between the two border routers in AS 213.

Consistent Route Selection Within the AS (Cont.)

Cisco.com



Q2: How will you influence the route selection on routers in AS 213 so that they select the fastest route?

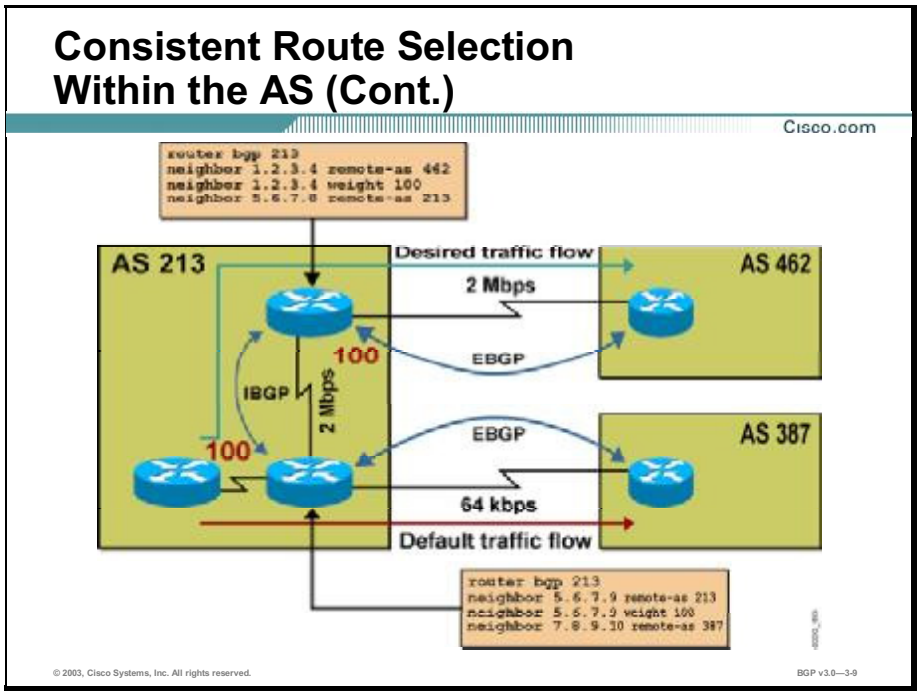
A2: By using weights on EBGP and IBGP sessions.

© 2003, Cisco Systems, Inc. All rights reserved.

One way of changing the default route selection is to use weights. Weight is an attribute that is locally significant to a router. It is a property or parameter and is, therefore, not seen on any neighboring routers. When designing BGP networks using weights, network administrators should set weights on every router. If there is more than one path for the same network, a router will choose the one with the highest weight. The default value for weight is 0.

In this example, the upper router in AS 213 sets a weight of 100 to routes received over the 2-Mbps link from AS 462 (primary link) and prefers them to possible internal updates from the bottom router, where the default weight is zero. The bottom router sets a weight of 100 to internal routes received from the upper router and prefers them to routes received from AS 387. As a result, all packets will leave the AS through the primary 2-Mbps link.

 Copyright © 2003, Cisco Systems, Inc.



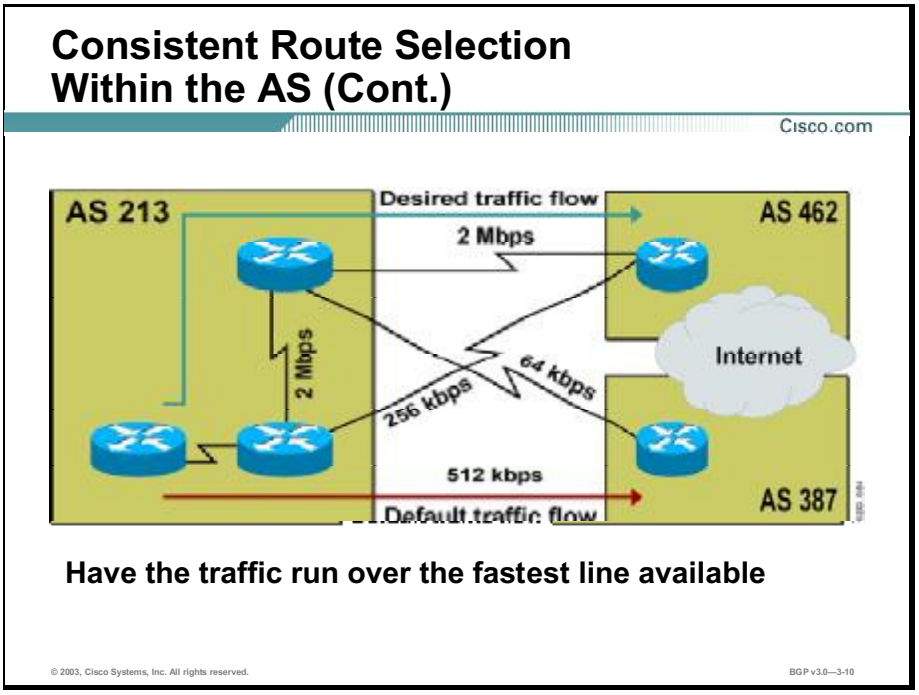
The configurations shown in the figure demonstrate how to change the default weight on a per-neighbor basis. If you use the **neighbor weight** command, all newly arrived updates will have a weight of 100. Updates coming from the other neighbor will still have the default weight of 0.

After you have applied the **neighbor weight** command, a refresh is needed from the neighbor. There are three ways of doing this, depending on the Cisco IOS version:

- Use **clear ip bgp neighbor address** to clear the neighbor relationship and re-establish it to refresh the BGP entries and apply the weight.
- Configure soft reconfiguration for the neighbor and use the **clear** command as shown in the figure. You can perform all subsequent clearing by using **clear ip bgp neighbor address soft in**, which does not reset the neighbor relationship. The soft reconfiguration feature is supported by Cisco starting with IOS version 11.2.
- Use **clear ip bgp neighbor address in** if both neighboring routers support the route refresh. The route refresh feature is available from Cisco starting with IOS version 12.1.

See the Implementing Changes in BGP Policy lesson for a detailed description of the commands here.

Example



This example is more complex. When you are trying to implement this example with weights, it requires two route-maps on each router within AS 213. Luckily, BGP has a similar mechanism that you can use for consistent AS-wide route selection: local preference.

Practice

Q1) What is a key difference between the local preference and weight attributes?

A) Local preference is local to the route on which it is configured.

B) Local preference is local to the AS within which it has been configured.

C) Local preference is local to the BGP administrative domain.

D) Local preference is global to a BGP domain.

Q2) What is an appropriate BGP implementation for the weight attribute?

A) all large-scale BGP implementations requiring AS-wide path selection policies

B) simple routing policies in smaller networks

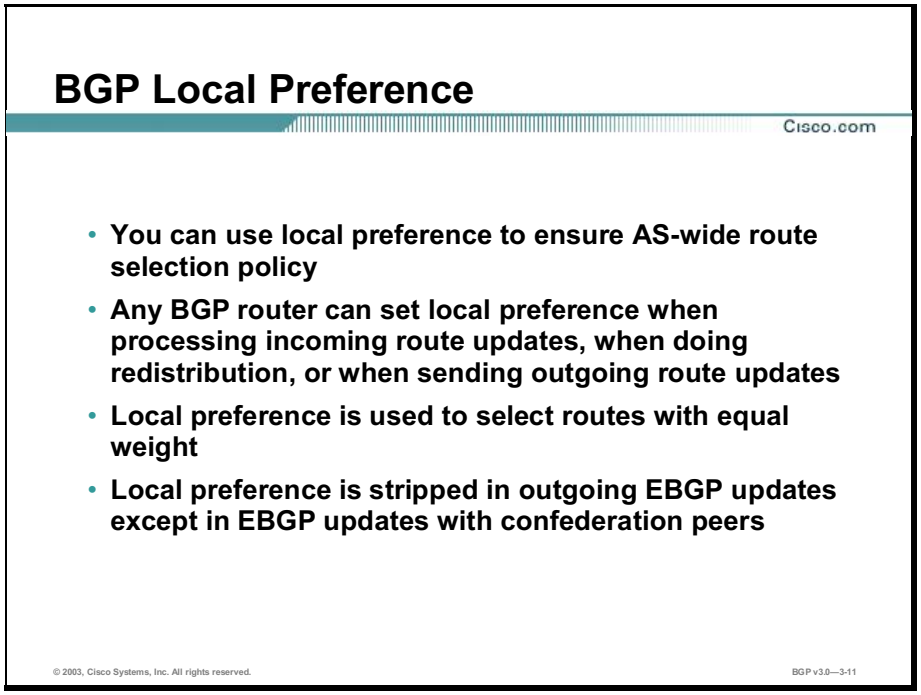
C) BGP implementations connecting a single-homed customer to provider

D) domains requiring global path selection policies

 Copyright © 2003, Cisco Systems, Inc.

BGP Local Preference

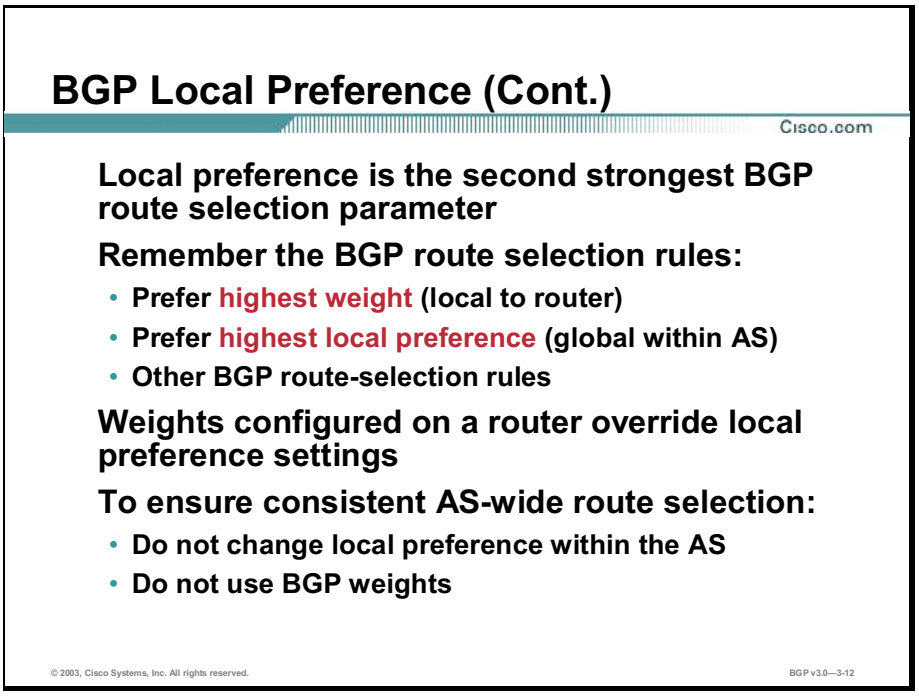
This topic describes how the BGP local preference attribute influences BGP route selection.


BGP Local Preference
Cisco.com
• You can use local preference to ensure AS-wide route selection policy
• Any BGP router can set local preference when processing incoming route updates, when doing redistribution, or when sending outgoing route updates
• Local preference is used to select routes with equal weight
• Local preference is stripped in outgoing EBGP updates except in EBGP updates with confederation peers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-11

Local preference is similar to weight; because it is as an attribute, you can set it once and then view it on neighboring routers without having to reset it. This attribute has a default value of 100, which the router will apply to locally originated routes and updates coming from external neighbors. Updates coming from internal neighbors already have the local preference attribute.

Local preference is the second strongest criteria in the route selection process. If there are two or more paths available for the same network, a router will first compare weight, and if the weights are equal for all paths, the router will then compare the local preference attribute. The path with the highest local preference value is preferred.

The local preference attribute is automatically stripped out of outgoing updates to External Border Gateway Protocol (EBGP) sessions. This fact means that you can use this attribute within a single AS only to influence the route selection process.


BGP Local Preference (Cont.)
Cisco.com
Local preference is the second strongest BGP route selection parameter
Remember the BGP route selection rules:
• Prefer highest weight (local to router)
• Prefer highest local preference (global within AS)
• Other BGP route-selection rules
Weights configured on a router override local preference settings
To ensure consistent AS-wide route selection:
• Do not change local preference within the AS
• Do not use BGP weights
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-12

Local preference is the second strongest BGP route selection parameter. Remember the route selection rules:

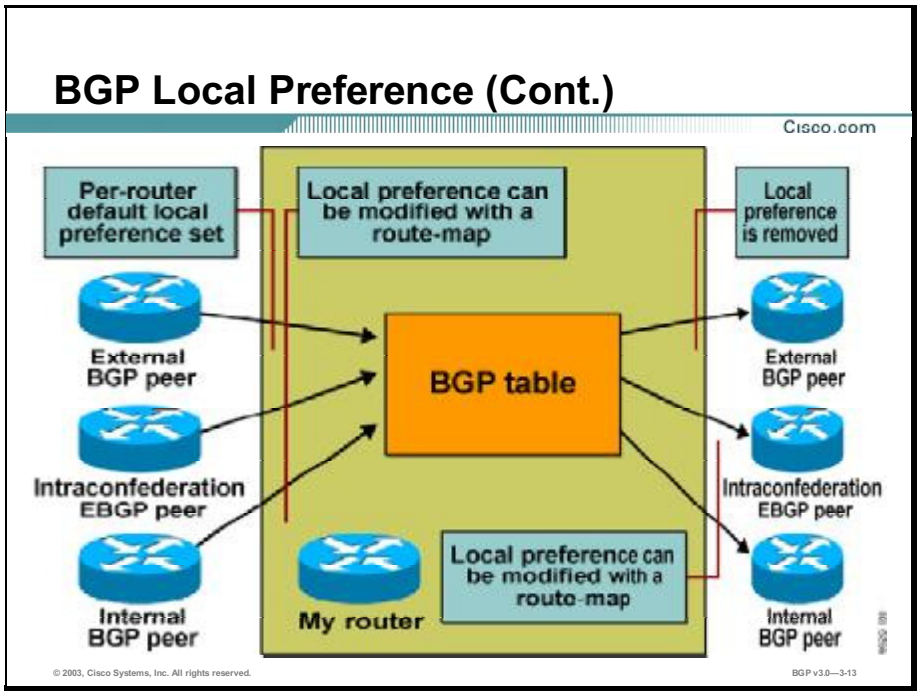
1. Prefer highest weight (local to router)
2. Prefer highest local preference (global within AS)
3. Process all remaining BGP route selection rules

Because network administrators can use both weight and local preference to manipulate the route selection process, they must decide which one to use. If local preference is used, the weight should be the same for all paths.

Network administrators can use weight on an individual router to override local preference settings used in the rest of the AS.

In most cases, it is enough to change the default local preference on updates coming from external neighbors. Network administrators should avoid changing the local preference attribute on internal sessions to prevent unnecessary complexity and unpredictable behavior.

 Copyright © 2003, Cisco Systems, Inc.



Network administrators can apply local preference in the following ways:

- Using a route-map with the **set local-preference** command. You can use the route-map on incoming updates from all neighbors or on outgoing updates to internal neighbors (not recommended).

- Using the **bgp default local-preference** command to change the default local preference value applied to all updates coming from external neighbors or originating locally.

Practice

Q1) If you configure both local preference and weight, which has the highest priority?

A) Local preference always has a higher priority.

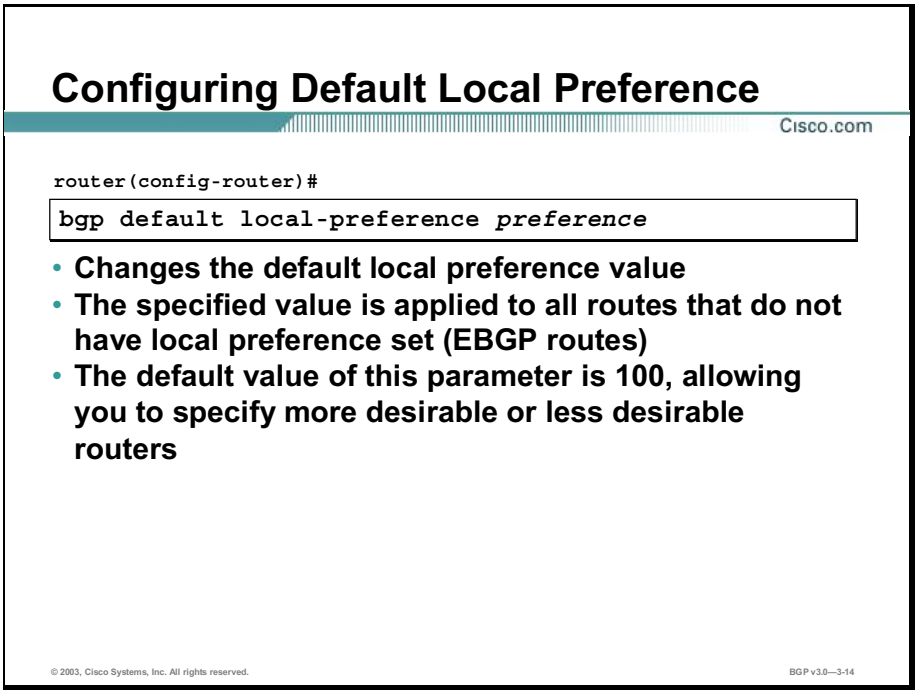
B) Weight always has a higher priority.

C) If routes are learned from different neighbors, the local preference will determine path selection.

D) If two routes have the same local preference, weight will determine the selected path.

Configuring Default Local Preference

This topic lists the Cisco IOS commands required to configure default BGP local preference on a Cisco router.


Configuring Default Local Preference
Cisco.com
router(config-router)#
bgp default local-preference preference
• Changes the default local preference value
• The specified value is applied to all routes that do not have local preference set (EBGP routes)
• The default value of this parameter is 100, allowing you to specify more desirable or less desirable routers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-14

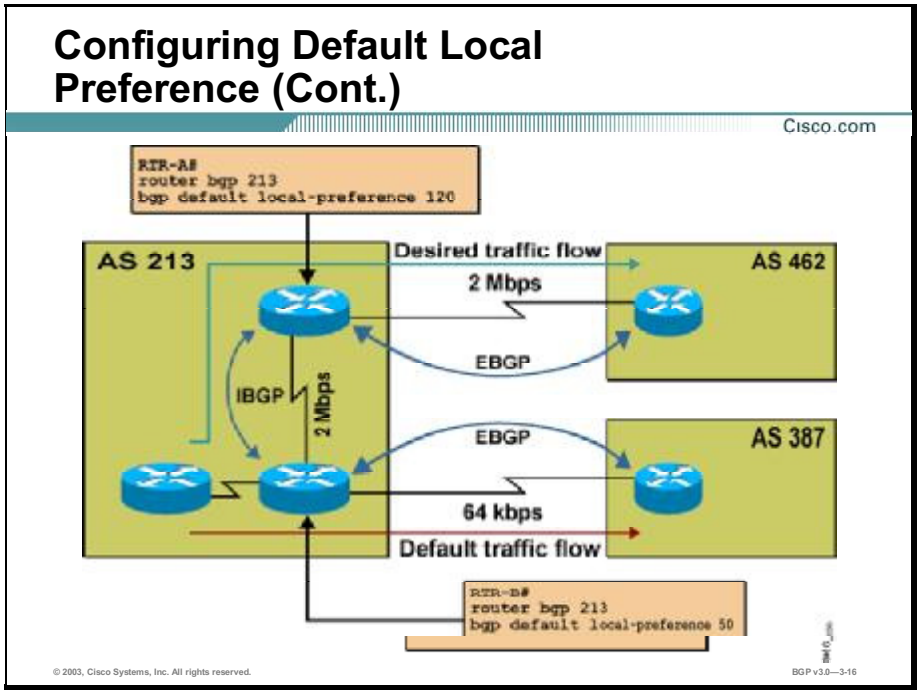
You can use the **bgp default local-preference** command in BGP configuration mode to change the default value of local preference. The new default value applies only to locally originated routes and those received from external neighbors.

Setting a value lower than the default of 100 will result in the router preferring internal paths to external (normally a router would prefer external routes).

Setting a value higher than 100 will result in external paths being preferred to all internal paths (also those with a shorter AS path).

 Copyright © 2003, Cisco Systems, Inc.

Example



In this example, the local preference attribute is used instead of weights. The two indicated routers in AS 213 have different default local preference values that are applied to external updates. The bottom router receives updates from the external neighbor and applies local preference to them. The same router then receives updates from the upper router, which set a local preference of 120 to all external updates. The bottom router then compares all paths and, where two paths exist, chooses the one with the higher local preference (120).

Practice

Q1) What is the default value of local preference?

A) 0

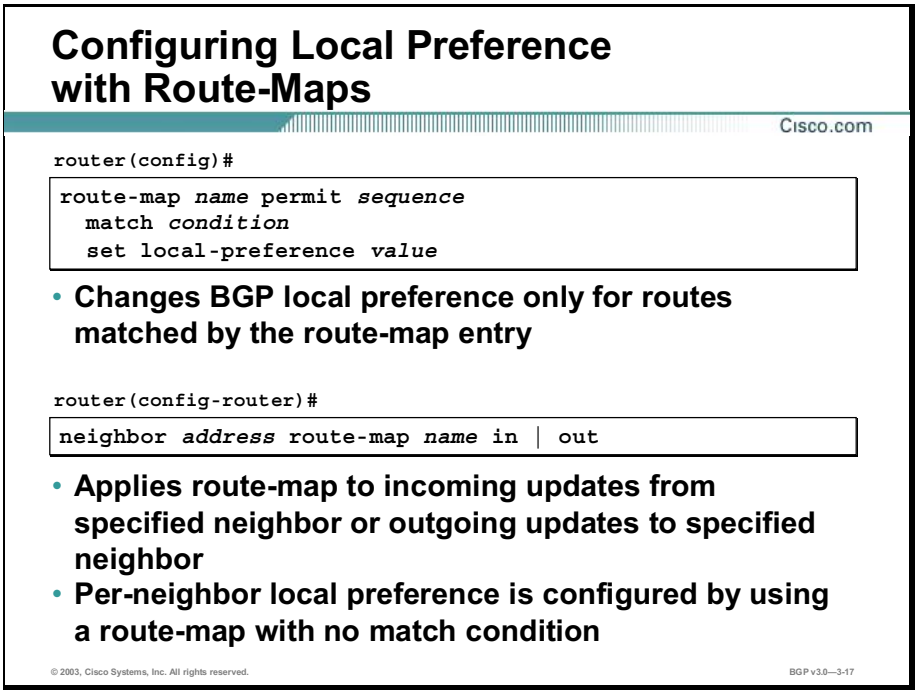
B) 100

C) 255

D) 32768

Configuring Local Preference with Route-Maps

This topic lists the Cisco IOS commands required to configure BGP local preference using route-map statements.


Configuring Local Preference with Route-Maps
Cisco.com
router(config)#
route-map name permit sequence
match condition
set local-preference value
• Changes BGP local preference only for routes matched by the route-map entry
router(config-router)#
neighbor address route-map name in | out
• Applies route-map to incoming updates from specified neighbor or outgoing updates to specified neighbor
• Per-neighbor local preference is configured by using a route-map with no match condition
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-17

To have more control over setting local preference, you may be forced to use a route-map. A route-map can have more statements, each with a different **set local-preference** command and a different match condition. If there is no **match** command, the route-map statement will apply local preference to all routes. The route-map can then be applied to BGP route updates in either the incoming or outgoing direction.

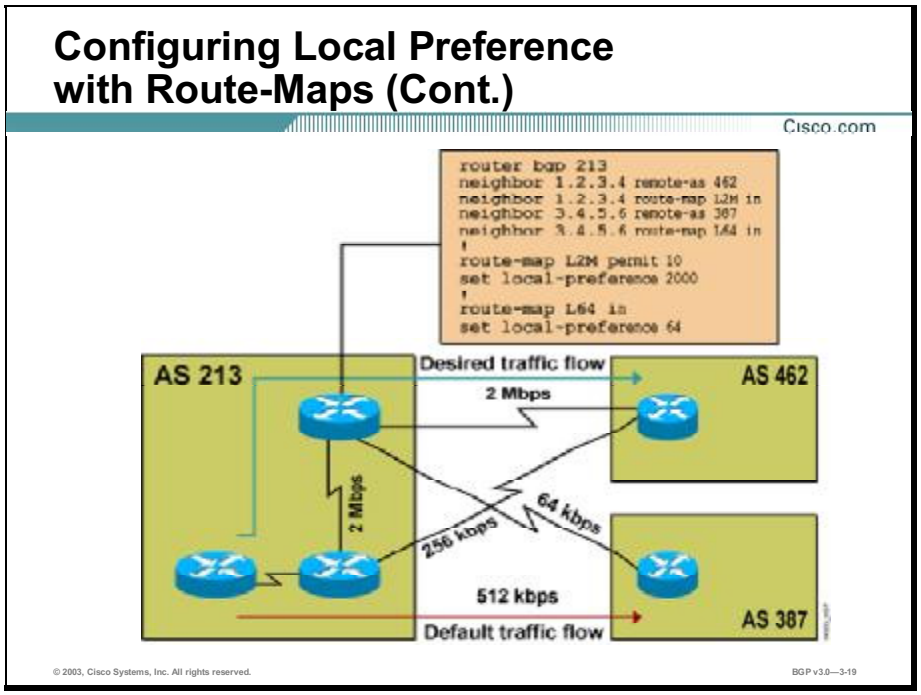
Note Applying a route-map to outgoing updates on external sessions will have no effect on local preference in the neighboring AS.

When routers use a route-map to set local–preference, the route-map is typically applied to incoming BGP routes advertised by an EBGP neighbor. The local router uses the local preference attribute in BGP route selection. In addition, the router also propagates the attribute to all Internal Border Gateway Protocol (IBGP) sessions in the local AS. Normally, no modifications of local preference are made on IBGP sessions. This restriction ensures that all routers in the local AS will use the same local-preference value and make the same decision in the route selection process.

Note If a network is not matched in any of the route-map statements, the network will be filtered. To permit unmatched networks without setting the local preference attribute, another route-map statement without **match** and **set** commands should be added to the end of the route-map. This statement should simply permit the remaining networks.

 Copyright © 2003, Cisco Systems, Inc.

Example



In this example, both routers have two external sessions. Using the **bgp default local-preference** command is no longer possible because the second fastest link is on another router.

The configuration here sets local preference according to the bandwidth of the link. A similar configuration exists on the bottom router. If the primary (2-Mbps) link fails, the paths learned through the bottom router in AS 213 (routes with a local preference of 512) will be used.

Practice

Q1) What effect does a route-map have on the local preference setting of outgoing EBGP updates?

A) The local preference can be set for outgoing EBGP updates to notify the neighboring AS about desired path selection.

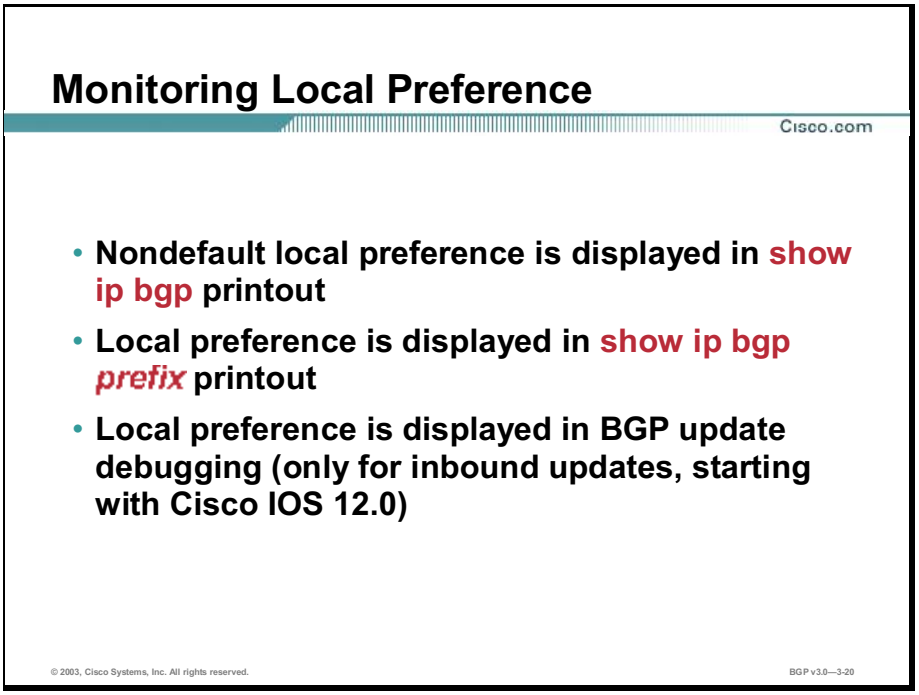
B) Route-maps cannot be used to apply changes to local preference in the outbound direction.

C) You can configure local preference only on IBGP neighbors.

D) Applying a route-map to outgoing updates on EBGP sessions does not affect local preference in the neighboring AS.

Monitoring Local Preference

This topic lists the Cisco IOS commands necessary to monitor BGP local preference.

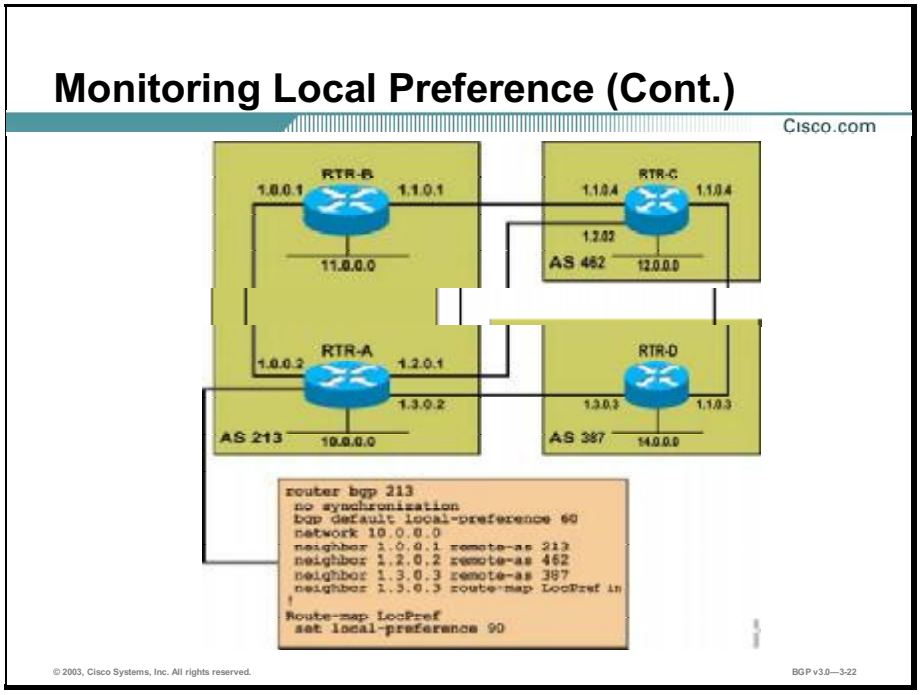

Monitoring Local Preference
Cisco.com
• Nondefault local preference is displayed in show ip bgp printout
• Local preference is displayed in show ip bgp prefix printout
• Local preference is displayed in BGP update debugging (only for inbound updates, starting with Cisco IOS 12.0)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-20

Although local preference is not a mandatory attribute, it is applied to every route. When you are using the **show ip bgp** command, a locally applied default value is not shown. All other values are displayed. You should use the command **show ip bgp** *prefix* to also display the locally applied value.

The output displayed from **show** and **debug** commands will vary depending on the Cisco IOS version. Newer versions typically display more information. In Cisco IOS version 12.0 and in later versions, enabling debugging of incoming routing updates will also display the local preference attribute.

 Copyright © 2003, Cisco Systems, Inc.

Example



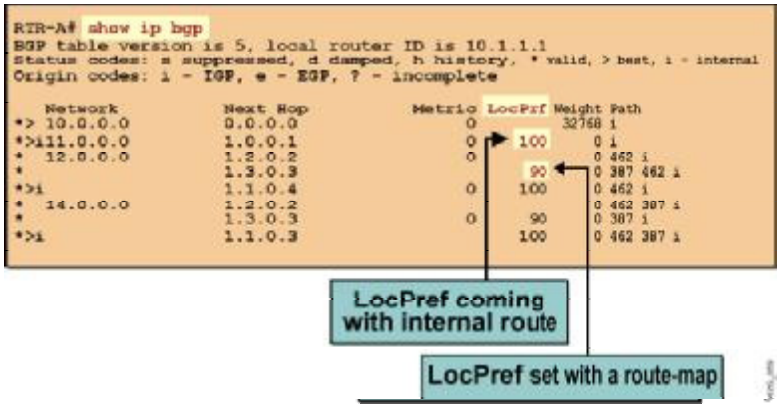
The network displayed in the figure was used to collect output from the **show** and **debug** commands in the next few examples. Every physical connection also includes a BGP session. All monitoring and troubleshooting commands were used on router RTR-A.

RTR-A has one internal and two external neighbors. RTR-B is setting local preference 100 to all updates, and RTR-A is setting a default local preference (value 60) for all external updates except for those coming from router RTR-D, where a route-map is used to set a local preference of 90. The following pages show the output of **show** and **debug** commands on router RTR-A.

Monitoring Local Preference (Cont.)

Cisco.com

Nondefault local preference is displayed in show ip bgp printout


RTR-A# show ip bgp
BGP table version is 5, local router ID is 10.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 10.0.0.0 0.0.0.0 0 32768 i
*>i11.0.0.0 1.0.0.1 0 100 0 i
* 12.0.0.0 1.2.0.2 0 0 462 i
* 1.3.0.3 90 0 387 462 i
*>i 1.1.0.4 0 100 0 462 i
* 14.0.0.0 1.2.0.2 0 462 387 i
* 1.3.0.3 0 90 0 387 i
*>i 1.1.0.3 100 0 462 387 i
LocPref coming with internal route
LocPref set with a route-map

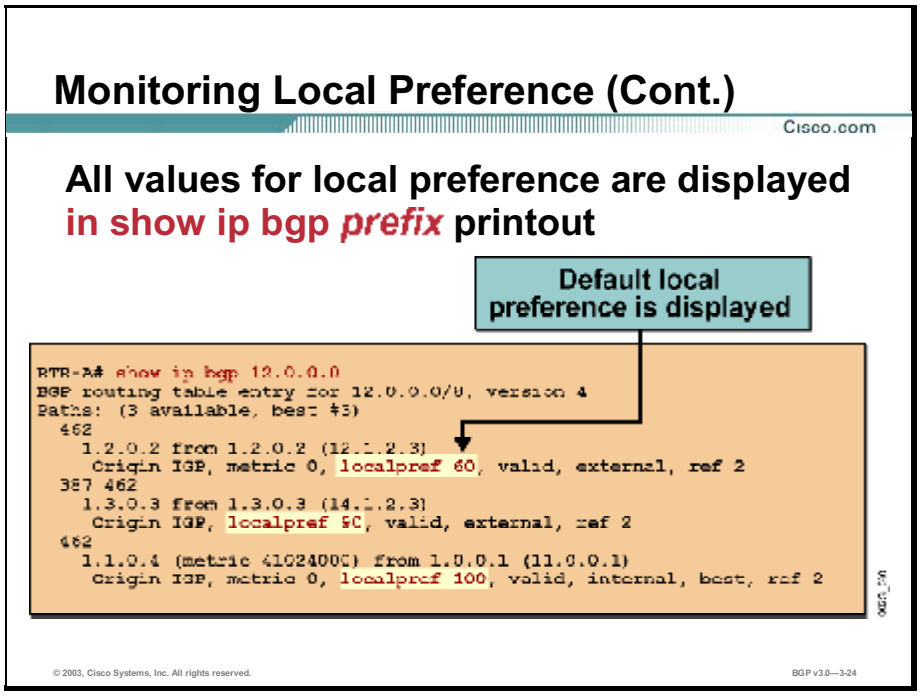
© 2003, Cisco Systems, Inc. All rights reserved.

The output shown in the figure contains routes with three different local preference values:

- Network 10.0.0.0/8 is locally originating on RTR-A, and the applied default local preference 60 is not displayed.
- The second path for network 12.0.0.0/8 was received from RTR-D and received a local preference value of 90 by the route-map.
- All routes received from router RTR-B are marked as internal and have a local preference value of 100 set on RTR-B.

Note The output of the **show ip bgp** command will not display the local preference value if the value is the same as the **bgp default local-preference** value in the local router. In the example, RTR-B was using its default local preference value (100). But when these routes propagated to RTR-A, then RTR-A displays the local preference value of 100 because it is different from the default local preference value configured on RTR-A.

 Copyright © 2003, Cisco Systems, Inc.

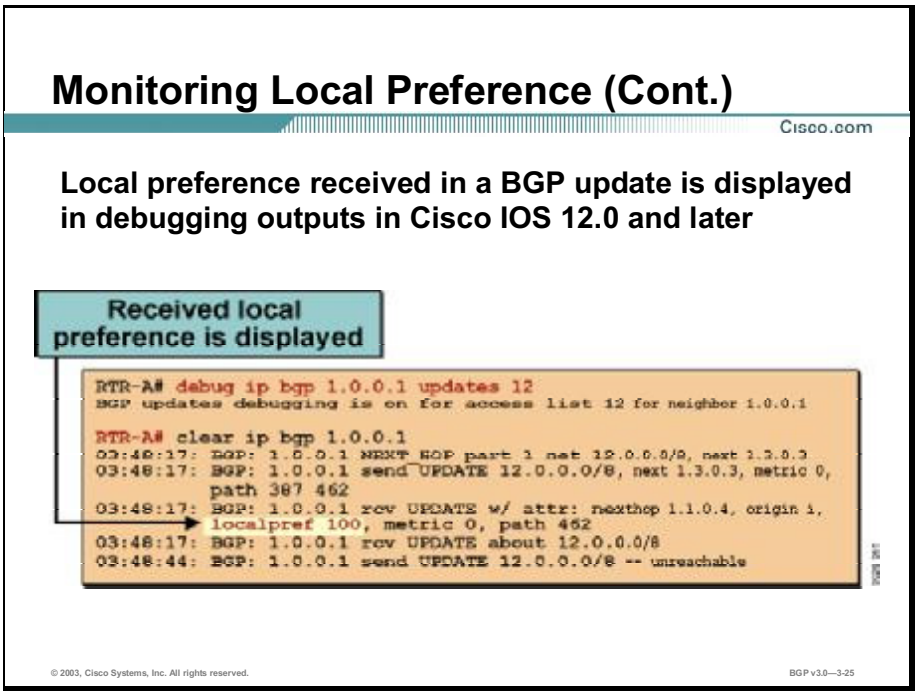

Monitoring Local Preference (Cont.)
Cisco.com
All values for local preference are displayed in show ip bgp prefix printout
Default local preference is displayed
RTR-A# show ip bgp 12.0.0.0
BGP routing table entry for 12.0.0.0/8, version 4
Paths: (3 available, best #3)
462
1.2.0.2 from 1.2.0.2 (12.1.2.3)
Origin IGP, metric 0, localpref 60, valid, external, ref 2
387 462
1.3.0.3 from 1.3.0.3 (14.1.2.3)
Origin IGP, localpref 90, valid, external, ref 2
462
1.1.0.4 (metric 41024000) from 1.0.0.1 (11.0.0.1)
Origin IGP, metric 0, localpref 100, valid, internal, best, ref 2
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-24

Use the **show ip bgp** *prefix* command to see more detailed information about a specific network, including the locally applied default local preference.

In this example, there are three different paths to reach the same network:

- The first path is external and was received from router RTR-C. The new default local preference value 60 was applied to the update.
- The second path is external and is received from router RTR-D. The route-map was used to set a local preference of 90.
- The third path is internal and was received from RTR-B. The update already contained a local preference attribute with a value of 100.

Router RTR-A chose the last path as best because it has the highest local preference.


Monitoring Local Preference (Cont.)
Cisco.com
Local preference received in a BGP update is displayed in debugging outputs in Cisco IOS 12.0 and later
Received local preference is displayed
RTR-A# debug ip bgp 1.0.0.1 updates 12
BGP updates debugging is on for access list 12 for neighbor 1.0.0.1
RTR-A# clear ip bgp 1.0.0.1
03:48:17: BGP: 1.0.0.1 NEXT_HOP part 1 net 12.0.0.0/8, next 1.3.0.3
03:48:17: BGP: 1.0.0.1 send UPDATE 12.0.0.0/8, next 1.3.0.3, metric 0, path 387 462
03:48:17: BGP: 1.0.0.1 rcv UPDATE w/ attr: nexthop 1.1.0.4, origin i, localpref 100, metric 0, path 462
03:48:17: BGP: 1.0.0.1 rcv UPDATE about 12.0.0.0/8
03:48:44: BGP: 1.0.0.1 send UPDATE 12.0.0.0/8 -- unreachable
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-25

This slide shows debugging output of incoming BGP updates. Since a router propagates the local preference attribute to other routers in the same AS only, local preference will be associated with routes sent from internal neighbors.

Practice

Q1) What command should you use to display locally applied local preference settings?

A) **show bgp preference detail**

B) **show ip bgp**

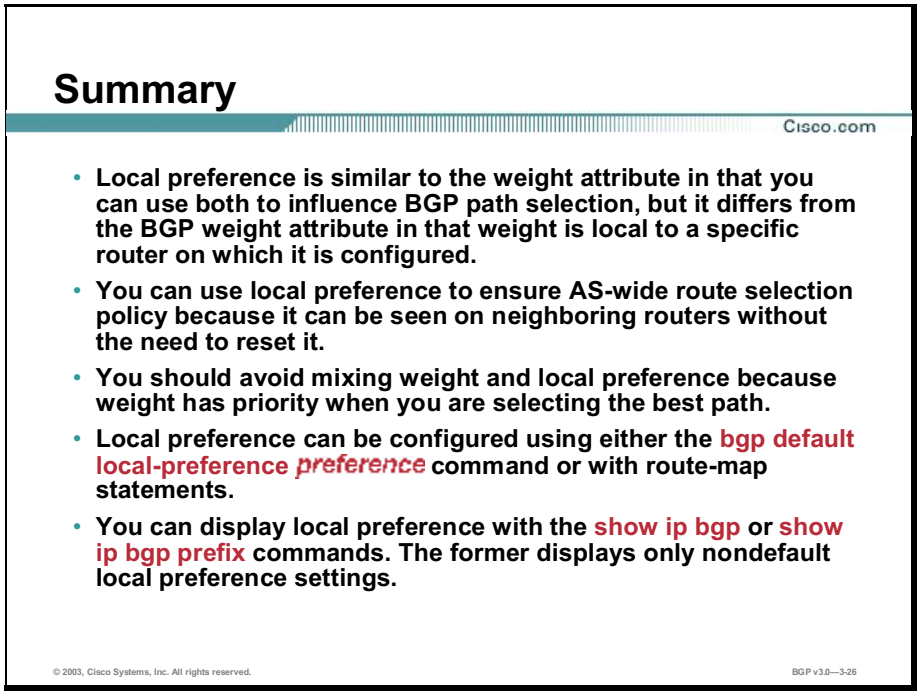
C) **show ip bgp detail**

D) **show ip bgp** *prefix*

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Local preference is similar to the weight attribute in that you can use both to influence BGP path selection, but it differs from the BGP weight attribute in that weight is local to a specific router on which it is configured.
• You can use local preference to ensure AS-wide route selection policy because it can be seen on neighboring routers without the need to reset it.
• You should avoid mixing weight and local preference because weight has priority when you are selecting the best path.
• Local preference can be configured using either the bgp default local-preference preference command or with route-map statements.
• You can display local preference with the show ip bgp or show ip bgp prefix commands. The former displays only nondefault local preference settings.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-26

Next Steps

After completing this lesson, go to:

- AS-Path Prepending lesson

References

For additional information, refer to these resources:

- For more information on BGP local preference, refer to “Border Gateway Protocol” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm
- For more information on BGP local preference, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL: http://www.cisco.com/warp/public/459/14.html

Laboratory Exercise: BGP Local Preference

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using the local preference attribute in a situation where you must support multiple connections to an Internet service provider (ISP).

After completing this exercise, you will be able to:

- Configure BGP local preference using route-maps
- Monitor BGP local preference

 Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
no synchronization	Disable BGP synchronization.
neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	Use this command to apply a route-map to incoming or outgoing routing updates.
set local-preference *num*	Use this command within a route-map to set the local preference attribute.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.
clear ip bgp	Restart the BGP session with your BGP neighbor.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- You want to establish two links with the "Good" service provider to increase the reliability of your Internet service. With several links connecting you to the same service provider, you must use local preference in your AS to ensure consistent AS-wide routing policy.

- In this exercise, you will establish the second link toward the "Good" service provider and use the local preference attribute to select the newly established link as the preferred exit point from your network.

- The additional link that you will establish connects WGxR2 and the "Good" router through the Frame Relay network. You will configure a BGP session (private peering) between WGxR2 and "Good" over this link. All traffic from your AS toward "Good" should flow over this link.

- On WGxR2, use data link connection identifier (DLCI) 20x and IP address 192.168.3x.1/30 for the Frame Relay link connection. The other side of this permanent virtual circuit (PVC) is connected to router "Good," which is already configured.

- The BGP routing design contains the following items:

  — AS x should prefer AS 20 as the upstream service provider. Router WGxR2 should be used as the exit point under normal circumstances.

  — Peering to AS 20 through WGxR1 should be used only if the primary link fails.

Copyright © 2003, Cisco Systems, Inc.

- Figure 1 shows the new physical connectivity, BGP sessions, and the expected traffic flow in the network.

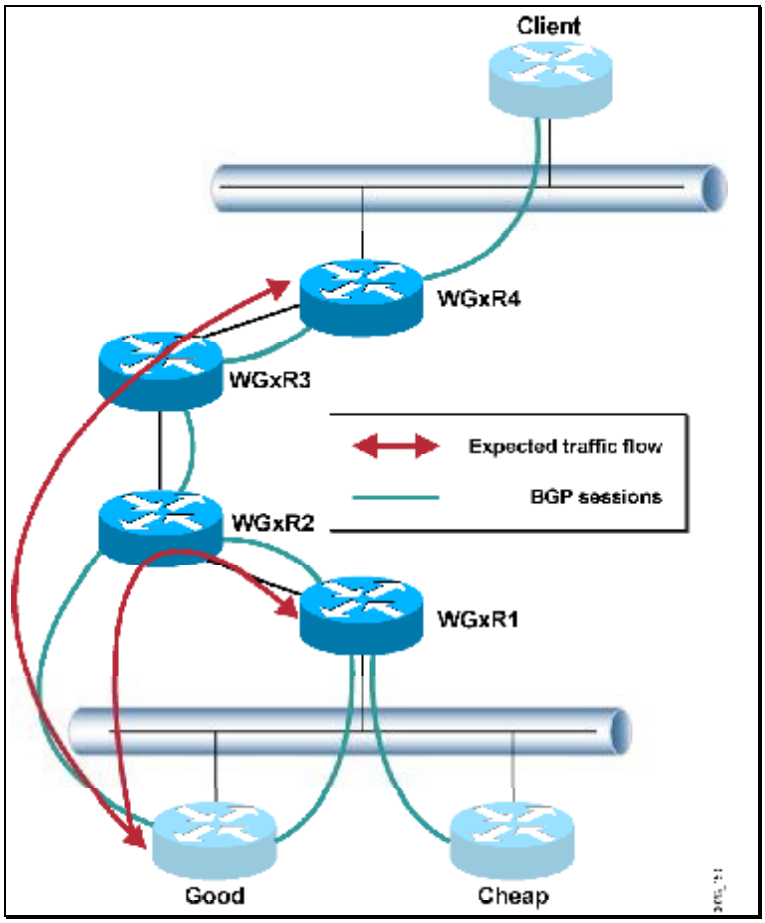


Figure 1: BGP local preference in a service provider network

Exercise Procedure

Complete these steps:

Configure additional IP connectivity:

Step 1 Create another point-to-point subinterface on WGxR2 using the parameters from the following table:

Router	IP address	DLCI
WGxR2	192.168.3x.1/30	20x

 Copyright © 2003, Cisco Systems, Inc.

Establish EBGP peering between WGxR2 and router “Good”:

Step 2 Configure the router “Good” as the BGP neighbor using the parameters from the following table:

Parameter	Value
Service provider IP address	192.168.3x.2
Service provider AS number	20

Establish IBGP peering between WGxR1 and router WGxR2:

Step 3 Establish an IBGP session between routers WGxR1 and WGxR2.

Step 4 Disable BGP synchronization on routers WGxR1 and WGxR2.

Use local preference to ensure AS-wide routing policy:

Step 5 On WGxR1, remove **neighbor weight** statements and any **route-map** statements that modify the weight attribute on WGxR1 for the routers “Good” and “Cheap.”

Step 6 Create a new route-map on router WGxR2. Within the route map, set the local preference higher than the default value of 100.

Step 7 Apply the route-map to incoming updates from router “Good.”

Step 8 Perform soft clearing of the BGP session by using the **clear ip bgp * [soft] in** command.

Exercise Verification

You have completed this exercise when you attain these results:

- On WGxR2, verify that the local preference has been changed and also note the IBGP routes.

```
wg1r2#sh ip bgp
BGP table version is 29, local router ID is 197.1.3.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.0.0.0         192.168.31.2             0    200      0 20 i
*> 128.20.0.0       192.168.31.2             0    200      0 20 i
*> 128.20.12.0/24   192.168.31.2             0    200      0 20 i
*> 128.22.0.0       192.168.31.2             0    200      0 20 i
*> 128.22.12.0/24   192.168.31.2                  200      0 20 22 i
*> 128.26.0.0       192.168.31.2             0    200      0 20 42 26
i
*> 128.37.0.0       192.168.31.2             0    200      0 20 42 37
i
*> 128.42.0.0       192.168.31.2             0    200      0 20 42 i
*> 128.51.0.0       192.168.31.2             0    200      0 20 42 26
51 i
*> 128.213.0.0      192.168.31.2             0    200      0 20 213 i
*> 128.214.0.0      192.168.31.2                  200      0 20 22
214 i
*> 192.20.11.0      192.168.31.2             0    200      0 20 i
*> 192.20.12.0/30   192.168.31.2             0    200      0 20 i
```

Copyright © 2003, Cisco Systems, Inc.

```
*> 192.22.11.0      192.168.31.2             0    200      0 20 i
*> 192.22.12.0/30   192.168.31.2                  200      0 20 22 i
*> 192.26.11.0      192.168.31.2             0    200      0 20 42 26
i
*> 192.37.11.0      192.168.31.2             0    200      0 20 42 37
i
*> 192.42.11.0      192.168.31.2             0    200      0 20 42 i
*> 192.51.11.0      192.168.31.2             0    200      0 20 42 26
51 i
* i192.168.1.0      192.168.1.1              0    100      0 i
*>                  0.0.0.0                  0         32768 i
*>i192.168.31.0     192.168.1.1              0    100      0 i
*> 192.213.11.0     192.168.31.2             0    200      0 20 213 i
*> 192.214.11.0     192.168.31.2                  200      0 20 22
214 i
*>i197.1.0.0/21     192.168.1.1                   100      0 i
*>i197.1.0.0/16     192.168.1.1                   100      0 i
*>i197.1.8.0/22     192.168.1.1                   100      0 i
*> 200.20.0.0/16    192.168.31.2             0    200      0 20 i
*> 200.22.0.0/16    192.168.31.2                  200      0 20 22 I
```

- Use traceroute from router WGxR4 and WGxR1 to 192.20.11.1.

```
WG1R4#traceroute 192.20.11.1
Type escape sequence to abort.
Tracing the route to 192.20.11.1

  1 192.168.1.9 32 msec 24 msec 20 msec
  2 192.168.1.5 32 msec 45 msec 40 msec
  3 192.168.31.2 56 msec *  52 msec

wg1r1#traceroute 192.20.11.1
Type escape sequence to abort.
Tracing the route to 192.20.11.1

  1 192.168.1.2 40 msec 28 msec 24 msec
  2 192.168.31.2 12 msec *  16 msec
```

- Use traceroute from router “Good” to interface loopback0 on routers WGxR4 and WGxR1.

```
Good#traceroute 197.1.7.1
Type escape sequence to abort.
Tracing the route to 197.1.7.1

  1 wg1 (192.168.20.1) 4 msec 0 msec 4 msec
  2 192.168.1.2 [AS 1] 20 msec 28 msec 24 msec
  3 192.168.1.6 [AS 1] 40 msec 45 msec 36 msec
  4 192.168.1.10 [AS 1] 52 msec *  48 msec

Good#traceroute 197.1.8.1
Type escape sequence to abort.
Tracing the route to 197.1.8.1

  1 wg1 (192.168.20.1) 4 msec *  0 msec
```

- Compare the two outputs of traceroute to determine if the routing is symmetrical. Both traceroute commands should show router WGxR2 in the path.

Answer these questions:

Q1) Is routing between router WGxR1 and “Good” symmetrical?

Q2) Which routers receive the local preference attribute?

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Prepending

Overview

In networks where connections to multiple providers are required, it is easy to set and control administrative policy for routes leaving an autonomous system (AS) under the same administrative control. A problem arises when administrative policies mandate a specific return path be used for traffic returning to the AS. In this case, administrative policy is difficult to control, because it requires service provider administrators to configure their routers to meet the customer administrative policy. One Border Gateway Protocol (BGP) mechanism that can potentially resolve the administrative policy issue is AS-path prepending. AS-path prepending potentially allows the customer to influence the route selection of its service providers.

This lesson introduces AS-path prepending by describing the problem that AS-path prepending solves. Also included in this lesson are the Cisco IOS® commands required to properly configure and monitor AS-path configurations. In addition, the lesson discusses special filtering requirements for upstream service providers connected to customers wishing to influence route selection using AS-path prepending.

Importance

When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. The “optimum” or “best” route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides many tools for administrators to influence route selection of both incoming and outgoing paths. One of these tools is AS-path prepending.

Objectives

Upon completing this lesson, you will be able to:

- Describe the need to influence BGP return path selection in a service provider environment
- Describe the function of AS-path prepending and how you can use it to facilitate proper return path selection

- Identify design considerations when you are implementing AS-path prepending to influence return path selection
- Identify the Cisco IOS commands required to configure AS-path prepending in a multihomed network
- Identify the Cisco IOS commands required to monitor the operation of AS-path prepending
- Describe concerns with using AS-path filters when neighboring autonomous systems require AS-path prepending

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module

Outline

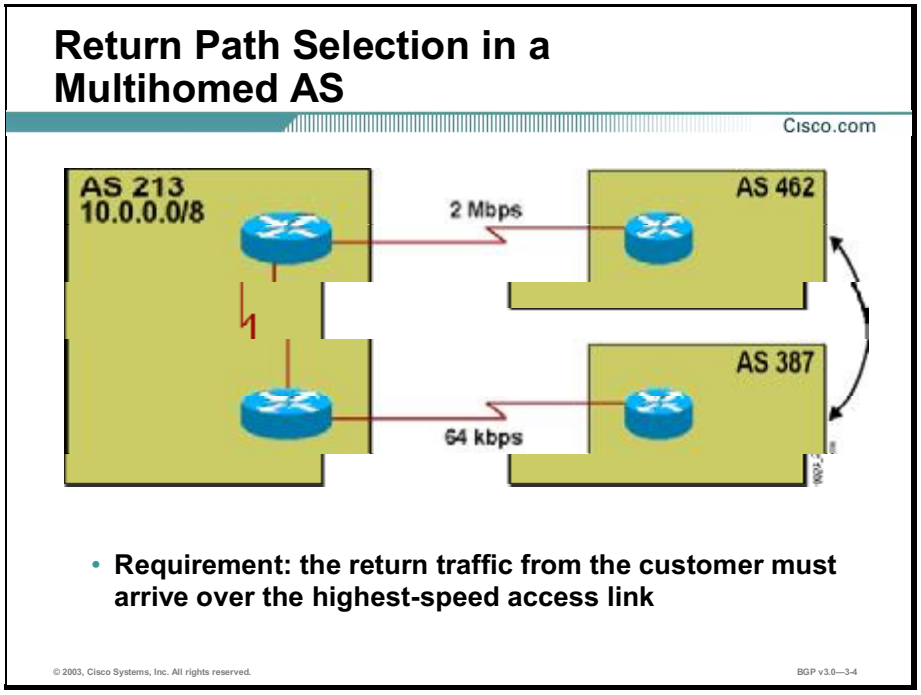
This lesson includes these topics:

- Overview
- Return Path Selection in a Multihomed AS
- AS-Path Prepending
- AS-Path Prepending Design Considerations
- Configuring AS-Path Prepending
- Monitoring AS-Path Prepending
- AS-Path Filtering Concerns with AS-Path Prepending
- Summary
- Assessment (Quiz): AS-Path Prepending

 Copyright © 2003, Cisco Systems, Inc.

Return Path Selection in a Multihomed AS

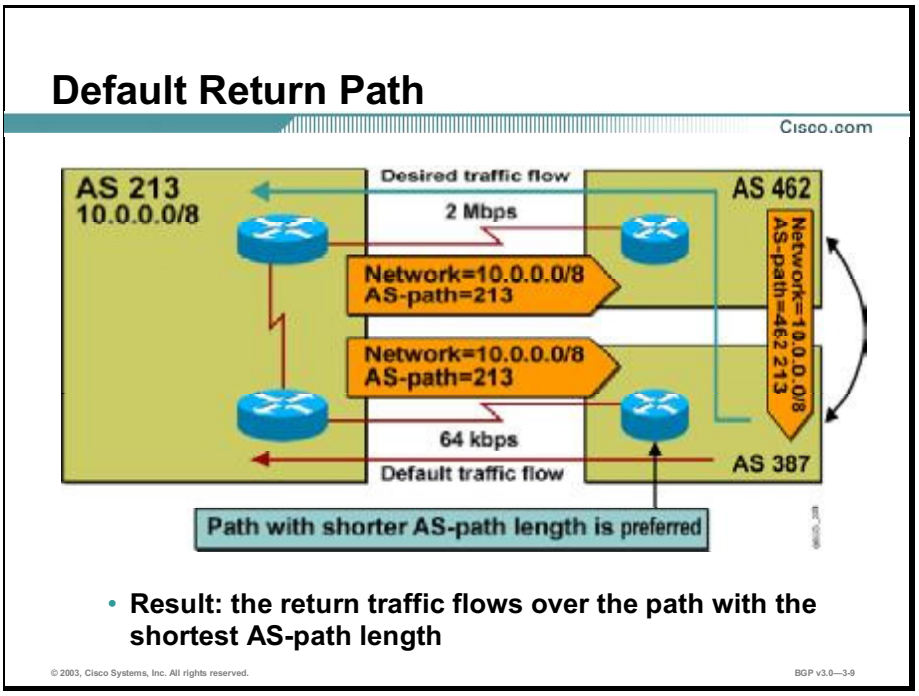
This topic describes the need to influence BGP return path selection in a service provider environment.


Return Path Selection in a Multihomed AS
Cisco.com
AS 213
10.0.0.0/8
2 Mbps
AS 462
AS 387
64 kbps
• Requirement: the return traffic from the customer must arrive over the highest-speed access link
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-4

It is fairly easy for an AS to select the appropriate path for outgoing traffic. It is much more complicated to influence other autonomous systems to select the appropriate path for traffic returning to a specific autonomous system.

To configure the preferred path only for outgoing traffic and not for incoming (return) traffic is likely to result in asymmetrical traffic flow as well as suboptimal performance of the return traffic. In the figure, outgoing traffic is directed to the high-speed line (2 Mbps) as a result of configuring local preference or weight. However, the return traffic from AS 387 would take the default path over the low-speed line (64 kbps). The low-speed line would be a limiting factor in the overall performance that the network could achieve.

In this example, AS 213 requests AS 387 to send packets toward network 10.0.0.0/8 via AS 462. The reason for this request is to improve network performance and minimize delay (assuming, of course, that the connectivity between AS 387 and AS 462 is better than the direct 64-kbps link between AS 387 and AS 213).



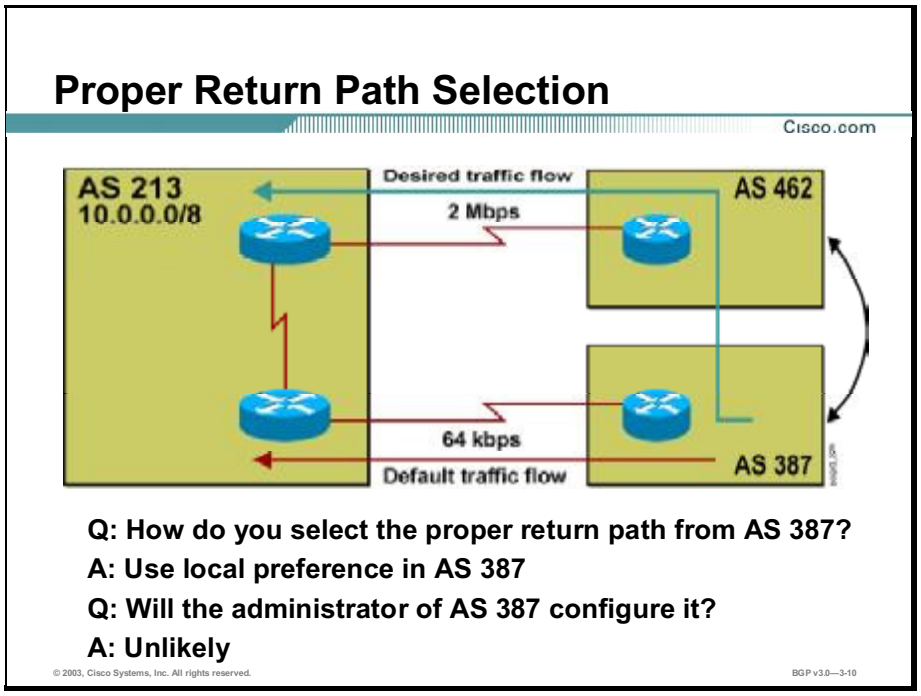
If no BGP path selection tools are configured on the route to influence the traffic flow, AS 387 will use the shortest AS-path. This action will result in unwanted behavior because the return traffic to AS 213 will be sent over the low-speed WAN link.

AS 213 announces network 10.0.0.0/8 over External Border Gateway Protocol (EBGP) sessions to both AS 462 and AS 387. When AS 213 sends EBGP updates, it changes the AS-path attribute according to BGP specifications. Both AS 462 and AS 387 receive a BGP update for network 10.0.0.0/8 with the AS path set to 213.

Because AS 462 selects the route for network 10.0.0.0/8 received from AS 213 as its best route, AS 462 will use that route and forward it on to AS 387. According to BGP specifications, AS 462 also changes the AS-path attribute. AS 387 receives the route to network 10.0.0.0/8 from AS 462 with an AS path set to 462 213.

AS 387 has now received two alternative routes to network 10.0.0.0/8 (the direct route from AS 213 and the route through AS 462). Because nothing is configured in AS 387 to influence the flow of traffic, the router will use the BGP route selection rule of shortest AS path to select the best return path to network 10.0.0.0/8.

 Copyright © 2003, Cisco Systems, Inc.



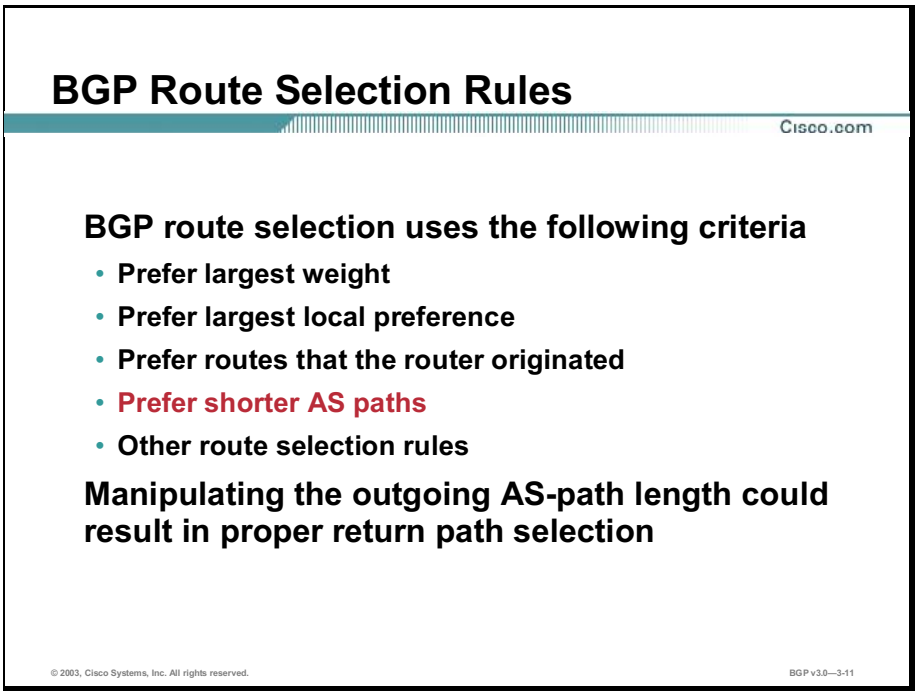
Remember that the incoming traffic flow (from the perspective of AS 213) will be a result of the route selection for outgoing traffic in AS 387. The traffic that is going out from AS 387 will end up as incoming traffic in AS 213.

If AS 387 configures some changes that cause the route selection process for outgoing traffic to prefer to reach network 10.0.0.0/8 via AS 462, it would result in behavior matching the desired administrative policy for AS 213, which specifies incoming traffic to the AS should be received over the high-speed link.

One way to accomplish the desired administrative policy in AS 213 is to configure the router in AS 387, which is receiving EBGP updates directly from AS 213, to assign a local preference value less than the default value (100) to all routes received from AS 213. The router in AS 387 is also configured specifically not to set local preference on EBGP routes received from AS 462. This results in assignment of the default value of 100 to all routes received from AS 462. When the route selection process in AS 387 selects the best route to reach network 10.0.0.0/8, the difference in local preference values causes AS 387 routers to select the path via AS 462 as the best.

However, all the configuration work to complete this process must be performed in AS 387. The network administrators of AS 387 would be required to modify the router configurations in AS 387 to satisfy the administrative policy requirements of AS 213. All changes must be documented and maintained according to the rules and procedures adopted by AS 387.

If AS 387 is a major Internet service provider (ISP), it is most likely that the network administrators are busy doing things other than tailoring router configurations based on the demand of a single leaf (nontransit) AS that lacks bandwidth on a redundant connection.


BGP Route Selection Rules
Cisco.com
BGP route selection uses the following criteria
• Prefer largest weight
• Prefer largest local preference
• Prefer routes that the router originated
• Prefer shorter AS paths
• Other route selection rules
Manipulating the outgoing AS-path length could result in proper return path selection
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-11

Recall that BGP route selection uses the following criteria:

1. Prefer largest weight
2. Prefer largest local preference
3. Prefer routes that the router originated
4. Prefer shorter AS paths
5. Then prefer all other route selection criteria

It is unlikely that the operator of an AS can request changes in router configurations in another AS. This situation makes it virtually impossible to influence another AS to select the desired path based on the weight and local preference attributes, because both would require configuration changes in the neighboring AS.

But if both the weight and the local preference parameters are left to their default settings, they will not indicate a difference. This situation causes the route selection process to continue down the list of selection criteria. The third criterion for selection will not influence route selection in this scenario, because none of the routes originated at the router performing the route selection. The fourth criterion will apply, however, because the AS paths have different lengths.

If the AS path is not manually manipulated by some administrative means, the path going over the fewest number of autonomous systems is selected by the router regardless of available bandwidth. However, if the AS that is attempting to influence the incoming traffic flow is sending out EBGP updates with a manipulated AS-path attribute over the undesired path, the receiver of this update is less likely to select it as the best because the AS-path now appears to be longer.

 Copyright © 2003, Cisco Systems, Inc.

The benefit of manipulating AS paths to influence the route selection is that the configuration needed is done in the AS that is requesting a desired return path.

Practice

Q1) What are two important reasons for influencing return path selection in a multihomed AS? (Choose two.)

A) to prevent asymmetrical traffic flow of traffic returning to the AS

B) to prevent network replay and denial-of-service attacks

C) to prevent suboptimal performance of return traffic that prefers to use lower-bandwidth links

D) to eliminate the need to influence outgoing path selection using BGP attributes

Q2) In a multihomed scenario, why does BGP not always use the highest-bandwidth link available?

A) because most administrators fail to configure interface bandwidths using the **bandwidth** command

B) because, ff BGP is not configured to properly redistribute into the IGP, bandwidth metrics are lost from their attached routes

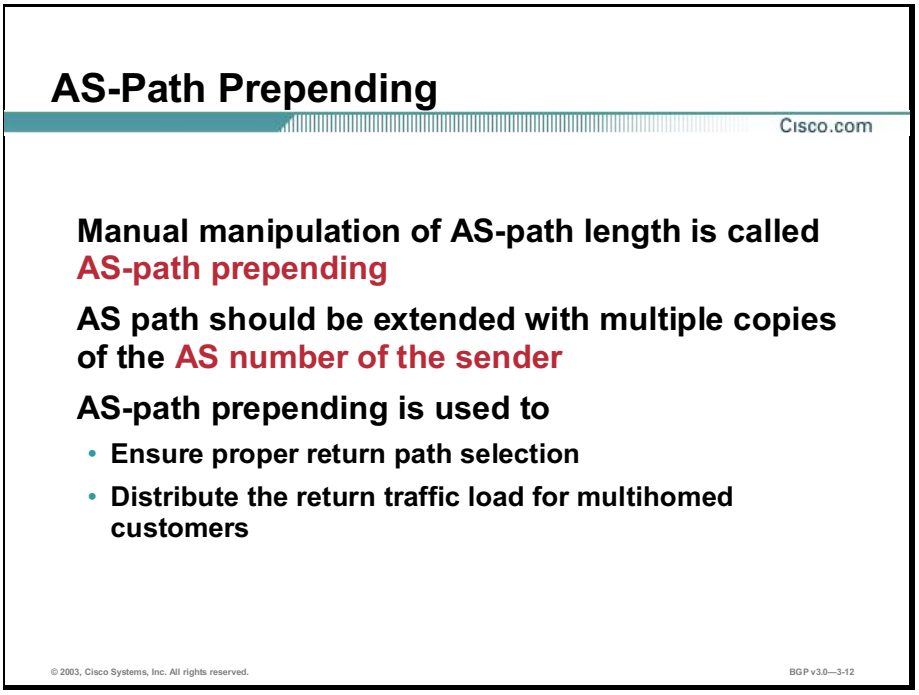
C) because many administrators fail to properly set the metric on each route as a reflection of the link bandwidth

D) because, if not influenced with attributes, the shortest AS path decides which path is used by BGP traffic

Copyright © 2003, Cisco Systems, Inc.

AS-Path Prepending

This topic describes the function of AS-path prepending and how you can use it to facilitate proper return path selection.


AS-Path Prepending
Cisco.com
Manual manipulation of AS-path length is called AS-path prepending
AS path should be extended with multiple copies of the AS number of the sender
AS-path prepending is used to
• Ensure proper return path selection
• Distribute the return traffic load for multihomed customers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-12

You can manipulate AS paths by prepending AS numbers to already existing AS paths. Normally, you perform AS-path prepending on outgoing EBGP updates over the nondesired return path. Because the AS paths sent out over the nondesired link become longer than the AS path sent out over the preferred path, the nondesired link is now less likely to be used as the return path.

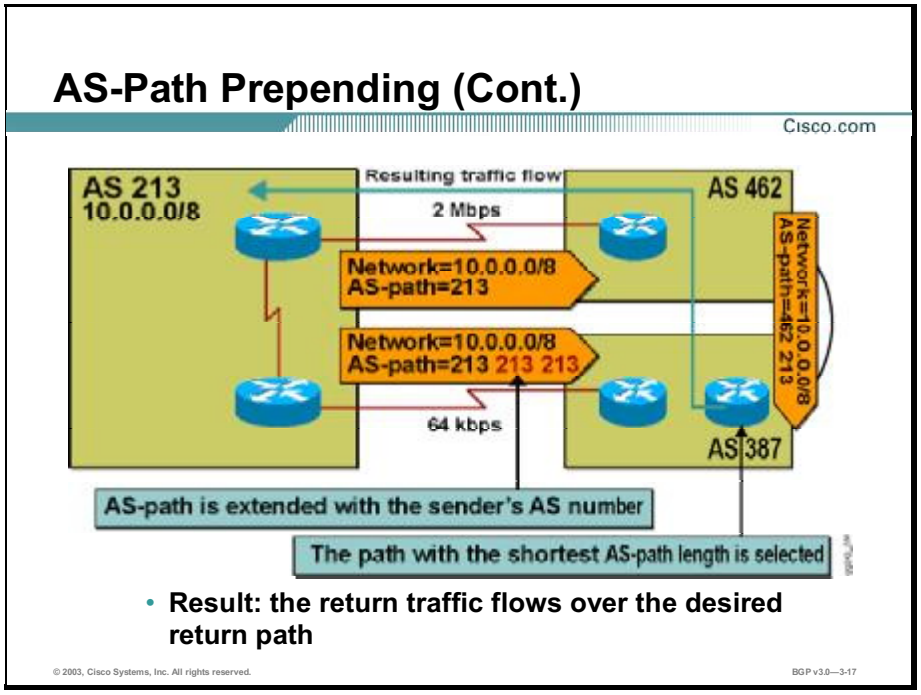
The length of the AS path is extended because additional copies of the AS number of the sender are prepended to (added to the beginning of) the AS-path attribute. To avoid clashes with BGP loop prevention mechanisms, no other AS number, except that of the sending AS, should be prepended to the AS-path attribute.

If another AS number is prepended in the AS path, the routers in the AS that has been prepended will reject the update due to BGP loop prevention mechanisms.

Prepending can be configured on a router for all routing updates sent to a neighbor or only on a subset of them.

 Copyright © 2003, Cisco Systems, Inc.

Example

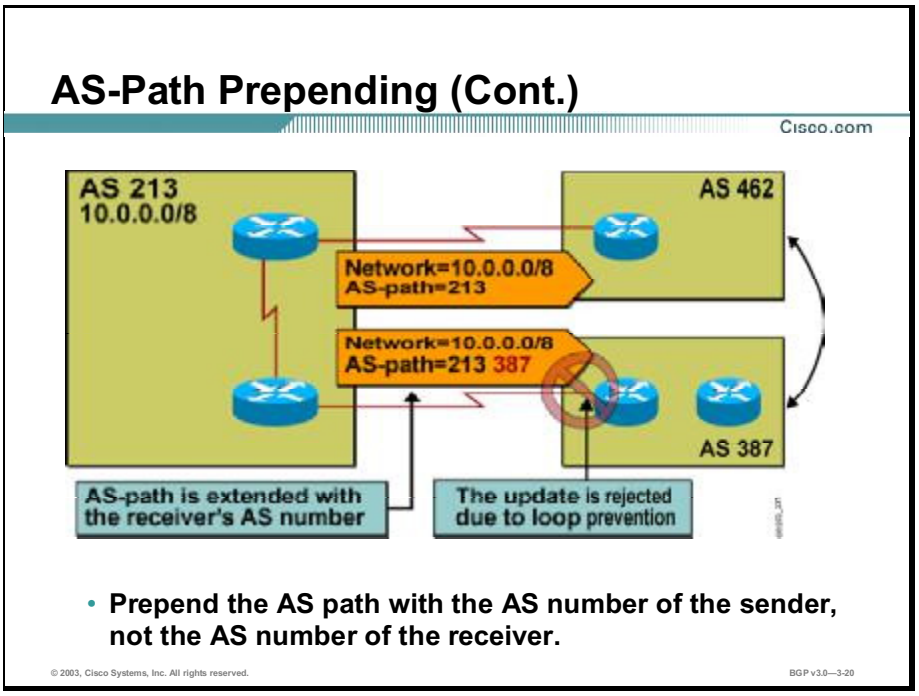


In this example, administrative policy in AS 213 prefers that the low-speed link be used for backup purposes only. As long as the high-speed link between AS 213 and AS 462 is available, all traffic should flow toward AS 213 using the high-speed link.

To accomplish this goal, you can configure the router in AS 213 that sends EBGP updates to AS 387 by prepending the AS path with two copies of the AS number 213. AS 387 receives two alternative routes to reach network 10.0.0.0/8: the update received directly from AS 213 (that has a manipulated AS path with a length of three) and the update received via AS 462 (that was not manually manipulated and therefore contains an AS-path length of two).

When AS 387 starts the route selection process to determine which route to use to reach network 10.0.0.0/8, it checks the AS-path length after the weight and local preference parameters. In this case, neither weight nor local preference has been configured, so the length of the AS path will be the deciding factor in the route selection process. Subsequently, AS 387 will prefer the shortest AS path and thus forwards packets toward network 10.0.0.0/8 via AS 462. The desired administrative policy will be met, and AS 213 will receive incoming traffic over the high-speed link.

If the forwarding path from AS 387 via AS 462 to AS 213 and network 10.0.0.0/8 is later broken, the BGP update to reach network 10.0.0.0/8 is revoked. In case of such a network failure, AS 387 will have only one remaining path to reach network 10.0.0.0/8. The route selection process has only one choice, the route directly to AS 213 over the low-speed WAN link. The low-speed link will therefore serve as backup to the high-speed WAN link.


AS-Path Prepending (Cont.)
Cisco.com
AS 213
10.0.0.0/8
AS 462
Network=10.0.0.0/8
AS-path=213
Network=10.0.0.0/8
AS-path=213 387
AS 387
AS-path is extended with the receiver's AS number
The update is rejected due to loop prevention
• Prepend the AS path with the AS number of the sender, not the AS number of the receiver.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-20

When manually manipulating AS-paths, the only valid AS number that you can prepend is the AS number of the sender. Prepending any other AS number will cause problems.

In the example, AS 213 is prepending AS number 387. The egress router performs AS-path prepending when the route is on its way to be transmitted to AS 387. After the manual manipulation is made, BGP automatically changes the AS path according to the BGP specifications. The local AS number should always be added first when updates are sent over an EBGP session. Therefore, when AS 387 receives the BGP update, the AS path contains the value 213 387. The AS number 387 was set there by the manual manipulation, and the AS number 213 was prepended automatically by BGP because the update was sent over an EBGP session.

When the edge router in AS 387 receives the BGP update, it checks the AS path to verify that the BGP updates were not propagated accidentally by a routing loop. Because the edge router finds its own AS number in the AS path, it assumes that the BGP update has already been in AS 387. According to the BGP specification, the update will be silently ignored.

Now assume that AS 213 had, for the manual manipulation, used yet a different AS number, not its own and not AS number 387. Would AS 387 now have accepted the update? The answer is yes. However, in this scenario, a problem would have appeared at a later stage when the route finally reached the actual AS belonging to the manually prepended AS number. This AS would have rejected the route because it would have found its own AS number somewhere in the AS path.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What are two functions of AS-path prepending? (Choose two.)

A) to mask the actual origination point of a route

B) to influence the path that a route takes out of its originating AS

C) to influence the path that a route takes into its originating AS

D) to distribute return traffic load between multihomed routers

Q2) What will happen if you prepend an AS number other than that of the originating AS to the AS path?

A) The upstream AS will automatically reject the route.

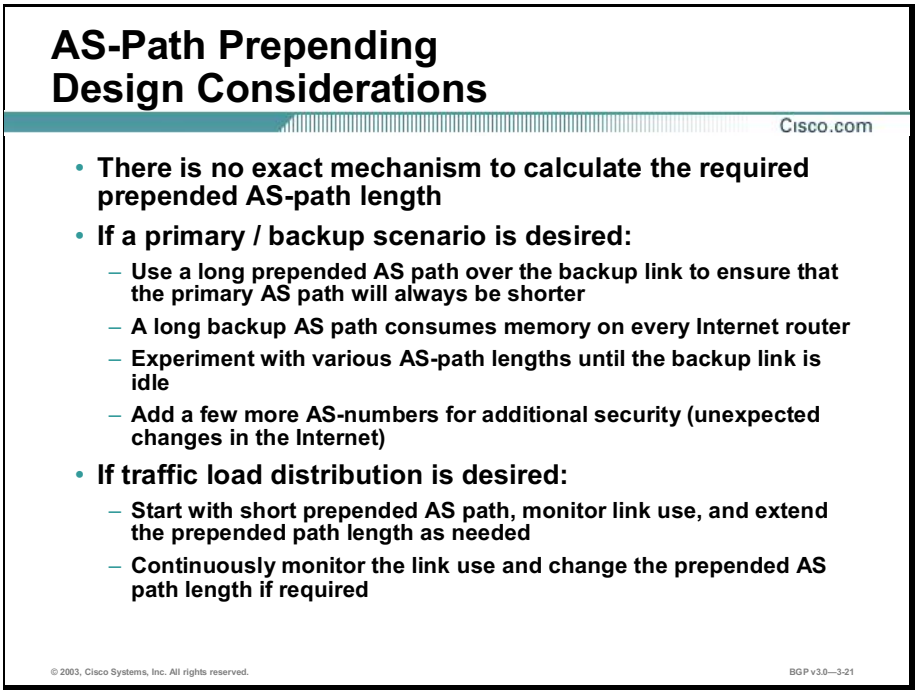
B) The route will traverse the Internet and be discarded by the downstream provider.

C) If the route is sent to the AS matching the prepended string, loop prevention mechanisms will silently ignore the update.

D) The AS-path prepending function will allow only the originating AS to be prepended to the AS path.

AS-Path Prepending Design Considerations

This topic identifies design considerations when you are implementing AS-path prepending to influence return path selection.


AS-Path Prepending
Design Considerations
Cisco.com
• There is no exact mechanism to calculate the required prepended AS-path length
• If a primary / backup scenario is desired:
– Use a long prepended AS path over the backup link to ensure that the primary AS path will always be shorter
– A long backup AS path consumes memory on every Internet router
– Experiment with various AS-path lengths until the backup link is idle
– Add a few more AS-numbers for additional security (unexpected changes in the Internet)
• If traffic load distribution is desired:
– Start with short prepended AS path, monitor link use, and extend the prepended path length as needed
– Continuously monitor the link use and change the prepended AS path length if required
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-21

How many copies of the AS number of the sender should you prepend to the AS path? The answer depends on the goals of the administrative policy. In the general case, it is not easy to determine the exact number of required AS numbers to prepend. The sending AS does not know what alternative paths are available to other autonomous systems.

Two typical cases where AS-path prepending is used for return path selection:

- Establishing a primary / backup link:

  As an announced backup (prepended) route propagates through the Internet, all the routers along the way that receive the route need to store it together with its AS-path attribute. If this information is long, it will consume extra memory in these routers. However, because routers forward only routes that are selected as “best,” an AS that receives multiple alternatives to a destination will select the route with the shortest AS path and forward only that route.

  In the case where both the primary and the secondary link are up, the neighboring AS will receive two routes to the same destination that differ only in the AS path length. The route with the shorter AS path will be subsequently advertised through the Internet.

  In the case where the primary link fails, the route with the longer AS path is the only remaining route. As a result, the primary route is withdrawn, and the prepended route is advertised through the Internet. In this case, extra memory will be consumed in each Internet router due to the storage of the prepended (longer) AS path.

 Copyright © 2003, Cisco Systems, Inc.

The longer the AS path announced to the EBGP neighbor on the other side of the backup link, the less likely it is that incoming traffic will be received from that neighbor. The network administrator can make a clever guess about how many copies of the AS number to prepend. After the prepending is implemented, the network administrator has to examine the result. If the expected result is not achieved, the configuration can be changed and a few more copies of the AS number can be prepended.

After AS-path prepending has successfully generated the desired results, the network administrator may take a step of precaution by prepending a few more copies of the AS number to the AS path. This action protects the customer from packets being routed over the backup link at a possible later stage when the topology between remote autonomous systems has unexpectedly changed, yielding a longer AS path to reach the primary link.

- Distributing the load of return traffic:

  In a multihomed scenario, there is no way to exactly predetermine the volume of traffic that will be received over a particular link. The traffic load on different links will change depending on where the senders are located (which autonomous systems they belong to). The network topology and the way that different remote autonomous systems are interconnected may also change with time, changing the load distribution. Only constant monitoring and fine-tuning will ensure that the desired results are achieved.

  In a first attempt at load distribution, the network administrator can configure a router connected to an overused link to prepend only a few extra copies of the local AS number. After the network has been given time to converge, the network administrator must check the change in load distribution. Monitoring of the load must be done for a period long enough to be statistically significant (several days or more). If enough volume of traffic has not moved from the overused link to the underused link, the administrator must prepend more copies of the local AS number, and the process of resending local routes and monitoring the results starts all over again.

Practice

Q1) What are two design principles for deciding how many AS number strings to prepend to an AS path? (Choose two.)

A) The AS-path length should be equal to one plus the number of AS-path hops to the destination, if load balancing is desired.

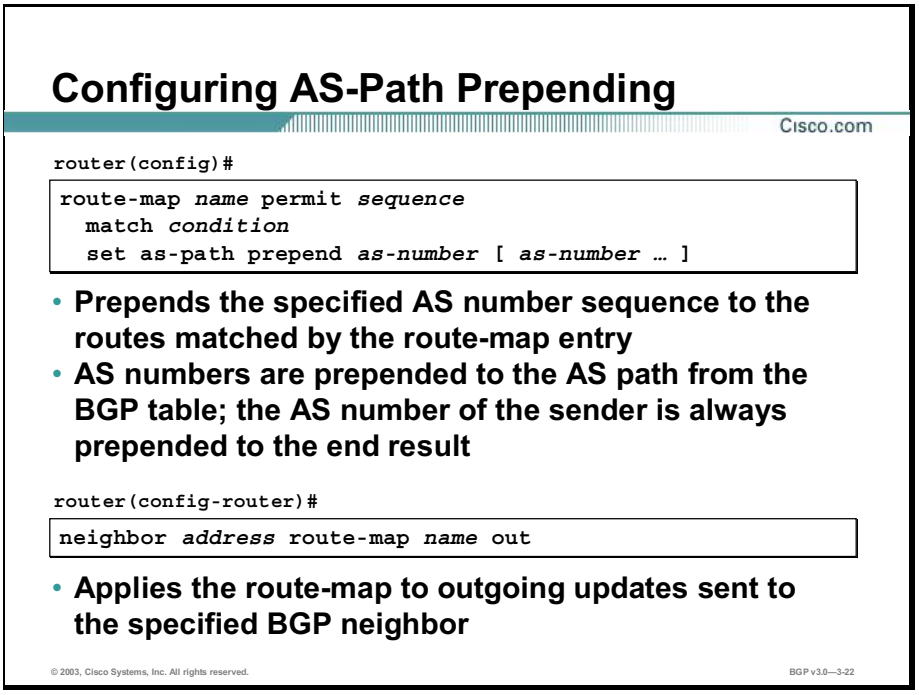
B) The number of AS numbers to append should equal the number of multihomed connections plus one.

C) If load distribution is desired, append as few AS numbers as possible to achieve the desired balancing goal.

D) If a backup link is desired, create a large AS path to ensure that the primary link is always used.

Configuring AS-Path Prepending

This topic identifies the Cisco IOS commands required to configure AS-path prepending in a multihomed network.


Configuring AS-Path Prepending
Cisco.com
router(config)#
route-map name permit sequence
match condition
set as-path prepend as-number [as-number …]
• Prepends the specified AS number sequence to the routes matched by the route-map entry
• AS numbers are prepended to the AS path from the BGP table; the AS number of the sender is always prepended to the end result
router(config-router)#
neighbor address route-map name out
• Applies the route-map to outgoing updates sent to the specified BGP neighbor
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-22

You can configure manual manipulation of the AS-path attribute (prepending) using a route-map with the **set as-path prepend** command. The route-map is used to prepend the specified AS numbers to outgoing EBGP route updates matched with the match condition, as specified in the route-map. AS-path prepending is completed first, and then the route is subject to the normal AS-path modification procedures when it is sent over an EBGP session.

You can also use the route-map to select only a subset of routes that should have their AS path manually manipulated. Use the **set as-path prepend** command with the appropriate **route-map permit** statement.

Note	Changing an outgoing route-map affects only the BGP updates sent after the change. In order to propagate the new and longer AS path with all announced routes, the routes must all be resent by the router. To do this, use the privileged EXEC command **clear ip bgp** with the **soft out** qualifier.

set as-path

To modify an AS path for BGP routes, use the **set as-path** route-map configuration command.

set as-path {**tag** | **prepend** *as-path-string*}

 Copyright © 2003, Cisco Systems, Inc.

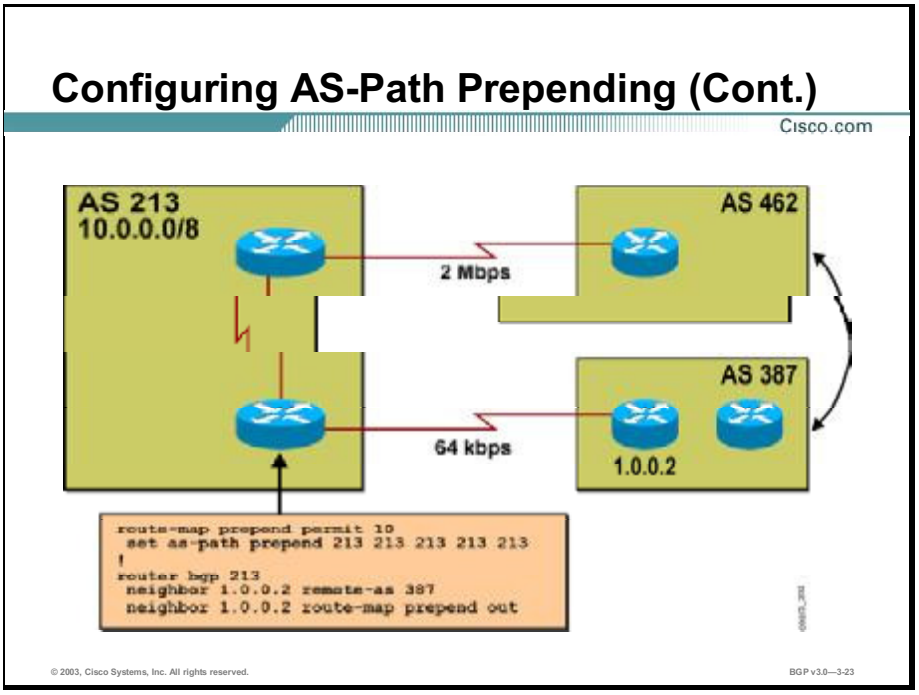
To not modify the AS path, use the **no** form of this command.

no set as-path {**tag** | **prepend** *as-path-string*}

Syntax Description

tag	Converts the tag of a route into an AS path. Applies only when redistributing routes into BGP.
prepend *as-path-string*	Appends the string following the keyword **prepend** to the AS path of the route that is matched by the route-map. Applies to inbound and outbound BGP route-maps.

Example



In this example, the lower router in AS 213 is configured to prepend its own AS number five times for all updates sent to the EBGP neighbor 1.0.0.2 in AS 387. This configuration will result in AS 387 receiving a route to network 10.0.0.0/8 with an AS path containing the AS number 213 six times (213 213 213 213 213 213). Cisco IOS software automatically adds the sixth copy of the AS number as the route leaves AS 213 in accordance with BGP specifications.

The configuration of the AS 213 router is completed in two steps:

Step 1 First, a route-map named “prepend” is created. The route-map selects all BGP routes and prepends five copies of 213 to the existing AS-path attribute already attached to each route. The lack of match conditions in the route-map indicates that all routes are matched.

Step 2 The route-map is applied to all outgoing updates to the EBGP neighbor 1.0.0.2.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) When you are configuring AS-path prepending with route-maps, when do the changes to the AS path take effect for already announced routes?

A) The changes are applied immediately.

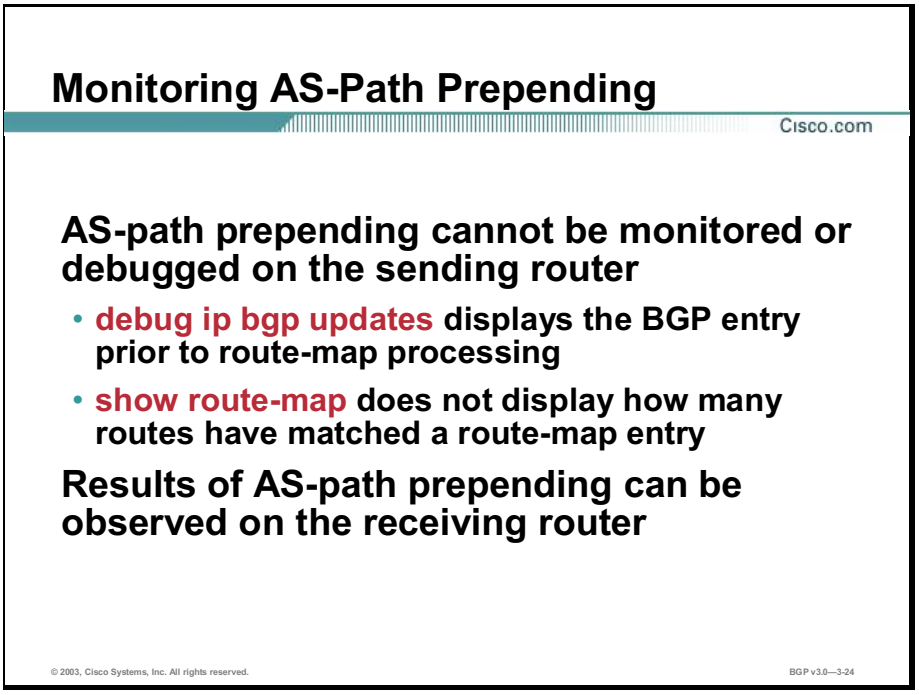
B) The changes will take effect upon clearing the BGP neighbor session of the configured neighbor.

C) The changes will take effect only after reloading the router.

D) Changes are applied only after the upstream AS accepts the AS path.

Monitoring AS-Path Prepending

This topic identifies the Cisco IOS commands required to monitor the operation of AS-path prepending.

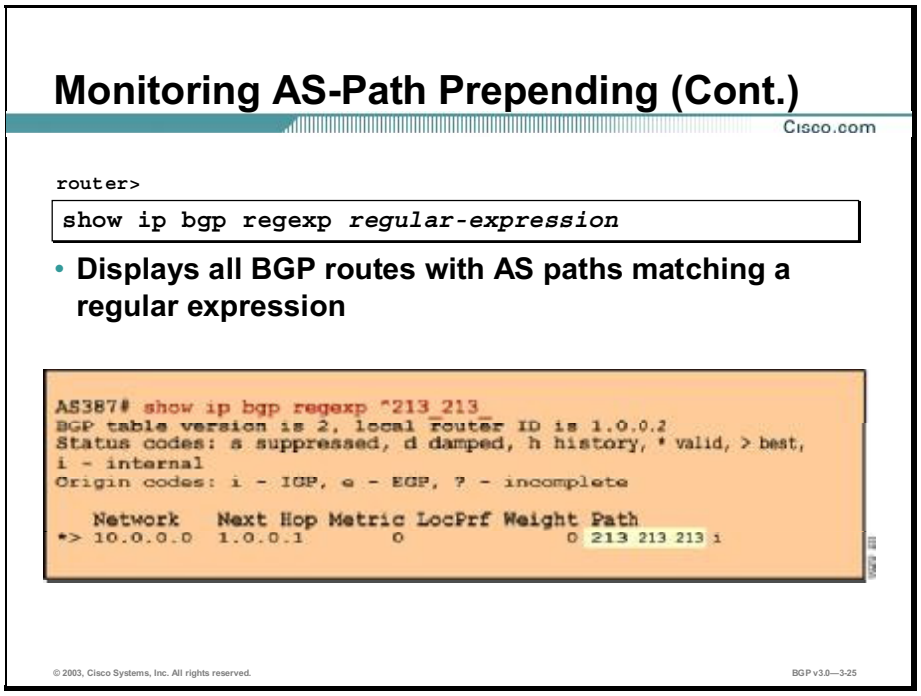

Monitoring AS-Path Prepending
Cisco.com
AS-path prepending cannot be monitored or debugged on the sending router
• debug ip bgp updates displays the BGP entry prior to route-map processing
• show route-map does not display how many routes have matched a route-map entry
Results of AS-path prepending can be observed on the receiving router
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-24

When you are monitoring AS-path prepending, the router doing the prepending is not the proper point to observe the results of the AS-path prepend operation. For instance, output from the **debug ip bgp updates** command does not display the prepended paths, because the route-map doing the prepending is applied afterward.

The **show route-map** command displays the configuration details of a route-map. The matching criteria and AS-path manipulation are displayed as output of the command. However, there is no indication of how many routes have been matched by a route-map statement and thus had their AS paths manipulated.

A better place for observing AS-path prepending is on the router receiving the BGP update containing the prepended AS path. At that point, you can use the pattern of AS number sequences in the received AS-path attribute of received routes to find the routes that have a prepended AS path.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring AS-Path Prepending (Cont.)
Cisco.com
router>
show ip bgp regexp regular-expression
• Displays all BGP routes with AS paths matching a regular expression
AS387# show ip bgp regexp ^213_213_
BGP table version is 2, local router ID is 1.0.0.2
Status codes: s suppressed, d damped, h history, * valid, > best,
i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 10.0.0.0 1.0.0.1 0 0 213 213 213 i
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-25

In the figure, the **show ip bgp regexp** command is used to find all the routes in the BGP table of the receiving router in AS 387 that are directly received from AS 213 and have at least one extra copy of AS number 213 in their AS path. Network 10.0.0.0/8 is displayed as the only route meeting this selection criterion. The AS path has been prepended with two extra copies of AS 213. The egress router in AS 213 automatically added the third copy of AS 213 because the route was sent across an EBGP session.

Practice

Q1) Where is the best place to monitor the AS-path prepending operation?

A) on the sending router using the **debug ip bgp updates** command

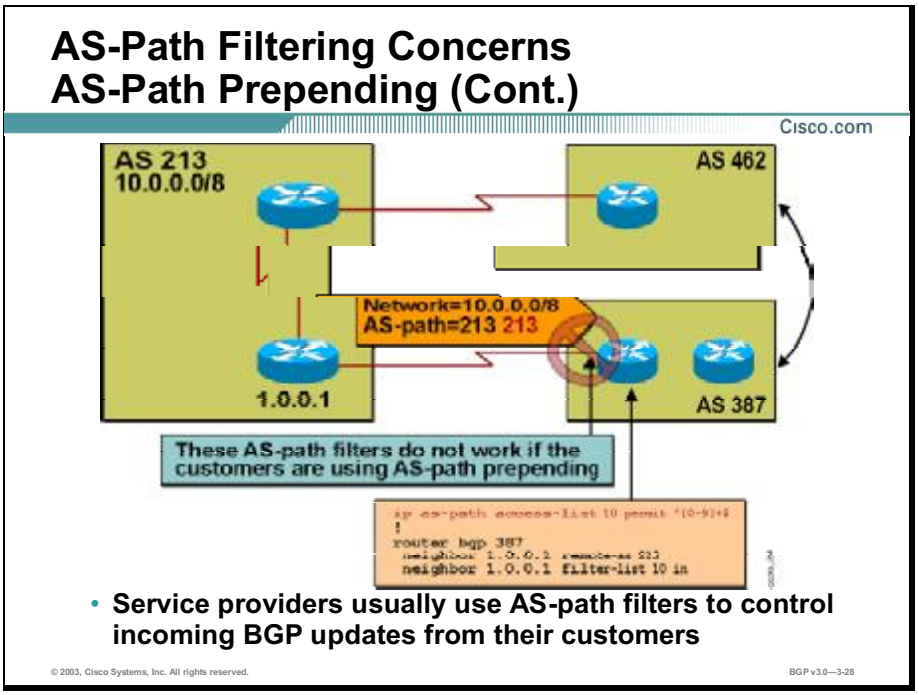
B) on the sending router using the **show route-map** command

C) on the receiving router using the **show ip bgp regexp** command

D) on either the sending or receiving router

AS-Path Filtering Concerns with AS-Path Prepending

This topic describes the concerns of using AS-path filters when neighboring autonomous systems require AS-path prepending.


AS-Path Filtering Concerns
AS-Path Prepending (Cont.)
Cisco.com
AS 213
10.0.0.0/8
AS 462
Network=10.0.0.0/8
AS-path=213 213
1.0.0.1
AS 387
These AS-path filters do not work if the customers are using AS-path prepending
router bgp 387
neighbor 1.0.0.1 filter-list 10 in
• Service providers usually use AS-path filters to control incoming BGP updates from their customers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-28

Service providers normally expect their customers to send routes originating only in the AS of the customer. However, because customers might not do so, proactive thinking and care for the rest of the Internet cause the service provider to implement AS-path filters on incoming updates received from their customers.

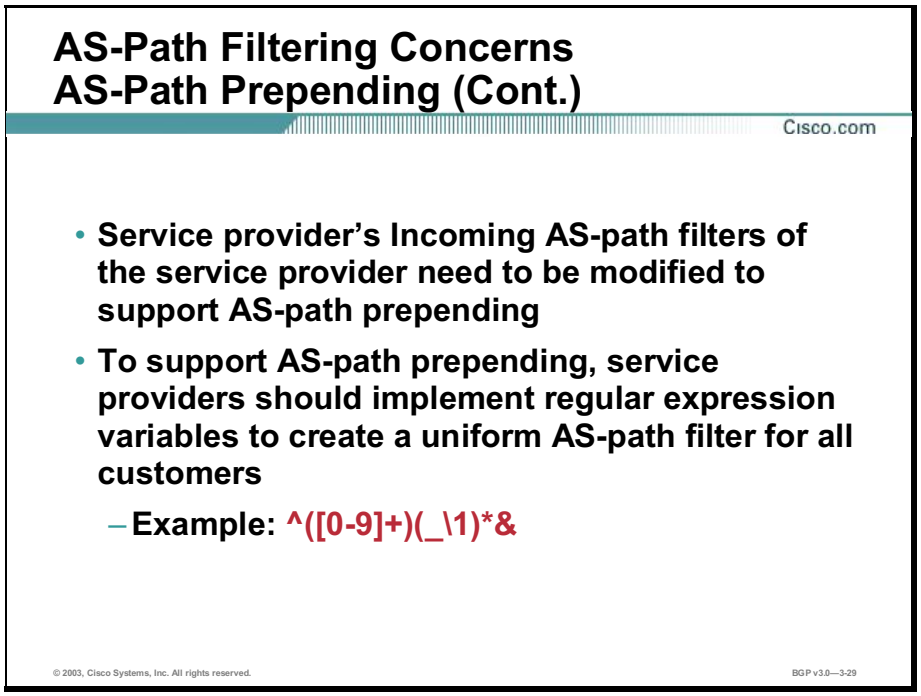
The network administrator of the service provider in the figure could configure individual filters for each neighbor. However, a single AS-path access-list permitting only AS paths with a length of exactly one AS number would be a better solution because the service provider can uniformly apply it to all incoming routes from all customers (possibly using a peer group).

In the figure, the service provider (AS 387) has configured a filter-list, which allows only AS paths that have a length of one AS number. When the customer changes its router configuration and starts to announce network 10.0.0.0/8 with a prepended AS path, the filter-list for incoming routes to AS 387 in the service provider router will filter those routes out. This filtering results in a situation where the network 10.0.0.0/8 is not reachable over the link between AS 213 and AS 387. Therefore, the backup function is not available.

Network 10.0.0.0/8 is, however, still reachable via the path going through AS 462. This situation means that AS 387 can send packets to network 10.0.0.0/8 but not over the direct link to AS 213. This failure may be hard to detect because, during normal conditions, all autonomous systems in the figure can exchange traffic.

 Copyright © 2003, Cisco Systems, Inc.

After AS 387 loses the route to network 10.0.0.0/8 via AS 462, possibly because the primary link between AS 213 and AS 462 is gone, the problem will be obvious. AS 387 can now no longer reach network 10.0.0.0/8 at all, although the physical link between AS 213 and AS 387 is available.


AS-Path Filtering Concerns
AS-Path Prepending (Cont.)
Cisco.com
• Service provider's Incoming AS-path filters of the service provider need to be modified to support AS-path prepending
• To support AS-path prepending, service providers should implement regular expression variables to create a uniform AS-path filter for all customers
– Example: ^([0-9]+)(_\1)*&
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-29

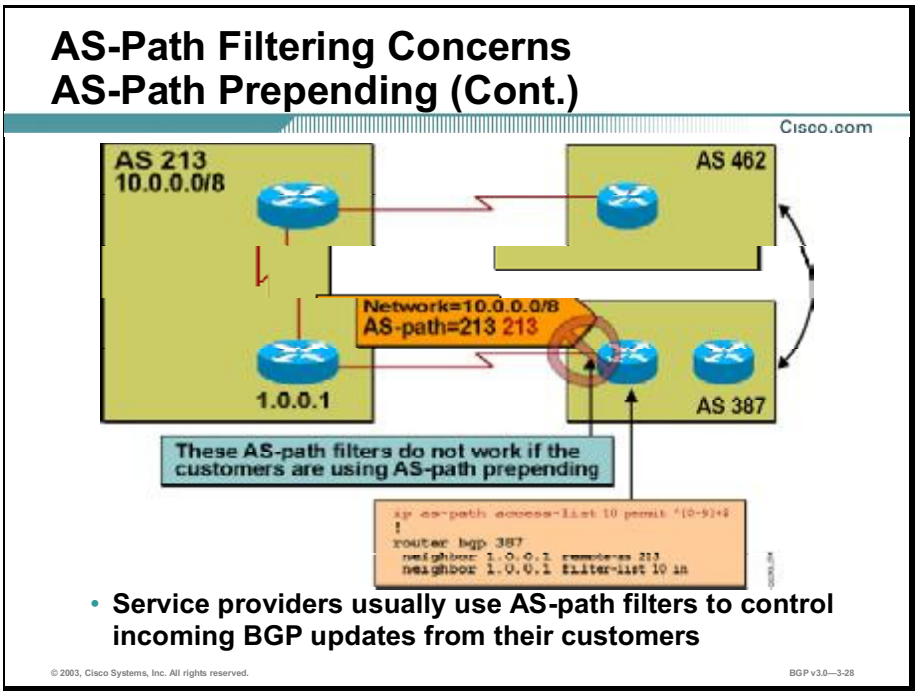
Because the AS of the service provider will receive customer routes with prepended AS-paths that have a length greater than one AS number, the provider must modify its incoming filters.

The service provider needs to create a new inbound regular expression filter, using regular expression variables and parentheses for recall.

What is needed is a filter that will allow any AS path containing one or multiple copies of the same AS number. An example of such a filter is:

- ^([0-9]+)(_\1)*$

This filter matches any AS path beginning with any AS number and continues with none or multiple repetitions of that same AS number (the variable "\1" repeats the value in the brackets). The regular expression would therefore match AS-paths 99 99 99, 2 2 2, or 100, but it would not match AS path 100 99.


AS-Path Filtering Concerns
AS-Path Prepending (Cont.)
Cisco.com
AS 213
10.0.0.0/8
AS 462
Network=10.0.0.0/8
AS-path=213 213
1.0.0.1
AS 387
These AS-path filters do not work if the customers are using AS-path prepending
router bgp 387
neighbor 1.0.0.1 filter-list 10 in
• Service providers usually use AS-path filters to control incoming BGP updates from their customers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-28

In the figure, the service provider (AS 387) has configured an individual filter for all routes received directly from AS 213. The AS path is required to start with 213. Then multiple copies of 213 may follow it. The asterisk allows for zero occurrences, permitting the AS path with a single copy of 213 as well.

If the same service provider router has more customers attached to it, they will all require an individual filter-list because the AS number of the customer is explicitly indicated in the regular expression.

An alternative would be to implement the AS-path filter using regular expression variables.

Practice

Q1) How may the AS of a service provider have to change its configurations when customers manipulate their outgoing AS paths?

A) Upstream providers have to disable all AS-path filtering.

B) The upstream provider must modify any filters based on the AS-path attribute to allow the prepended path.

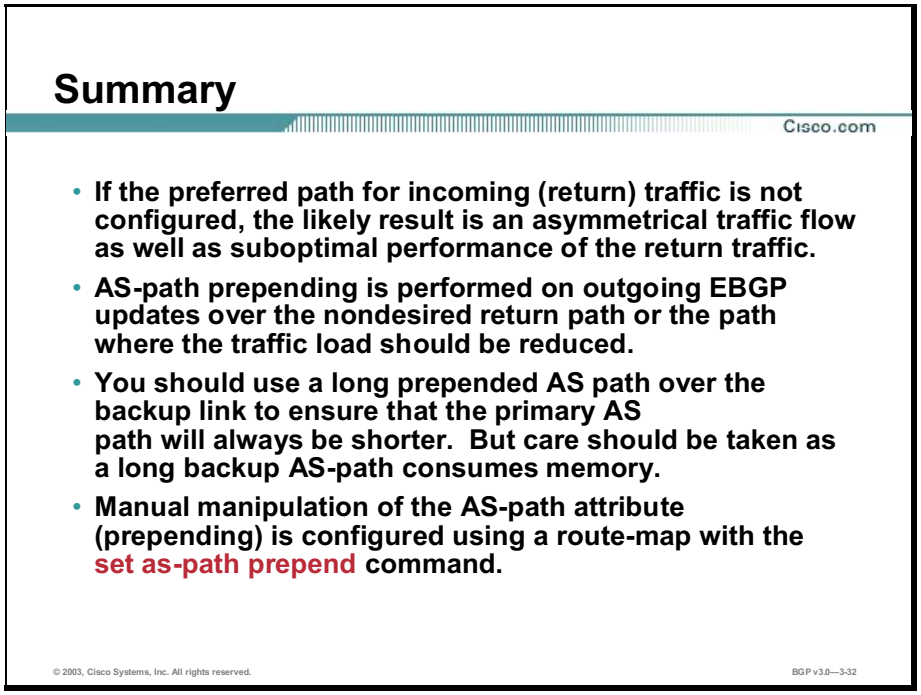
C) The provider has to create an AS-path filter specific to each customer.

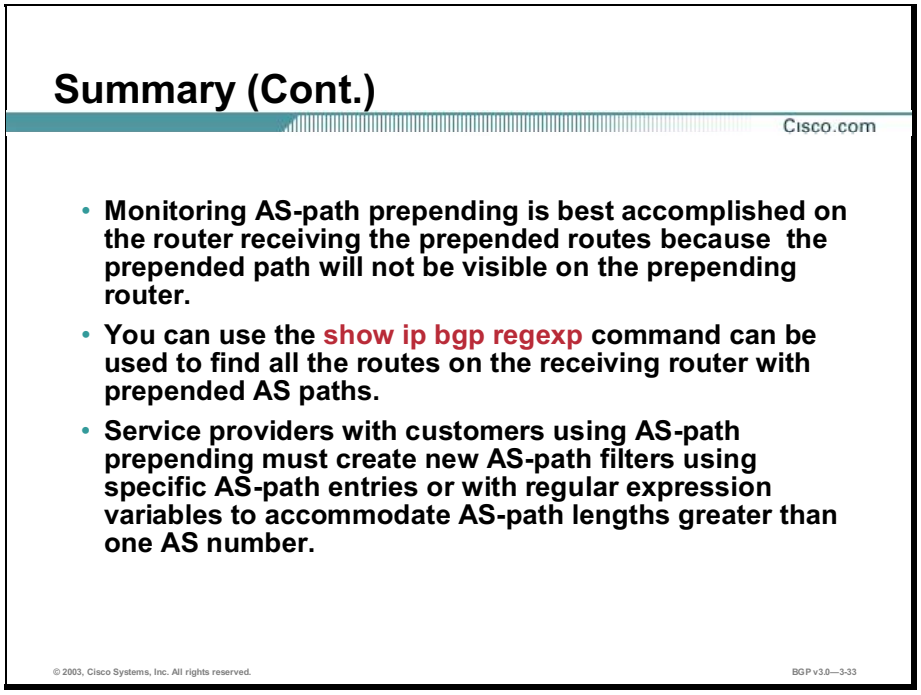
D) Customer-manipulated AS paths typically require no changes in the provider network.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• If the preferred path for incoming (return) traffic is not configured, the likely result is an asymmetrical traffic flow as well as suboptimal performance of the return traffic.
• AS-path prepending is performed on outgoing EBGP updates over the nondesired return path or the path where the traffic load should be reduced.
• You should use a long prepended AS path over the backup link to ensure that the primary AS path will always be shorter. But care should be taken as a long backup AS-path consumes memory.
• Manual manipulation of the AS-path attribute (prepending) is configured using a route-map with the set as-path prepend command.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-32


Summary (Cont.)
Cisco.com
• Monitoring AS-path prepending is best accomplished on the router receiving the prepended routes because the prepended path will not be visible on the prepending router.
• You can use the show ip bgp regexp command can be used to find all the routes on the receiving router with prepended AS paths.
• Service providers with customers using AS-path prepending must create new AS-path filters using specific AS-path entries or with regular expression variables to accommodate AS-path lengths greater than one AS number.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-33

Next Steps

After completing this lesson, go to:

- BGP Multi-Exit Discriminator lesson

References

For additional information, refer to these resources:

- For more information on AS-path prepending, refer to “BGP Case Studies Section 3” at the following URL: http://www.cisco.com/warp/public/459/15.html

 Copyright © 2003, Cisco Systems, Inc.

Quiz: AS-Path Prepending

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe the need to influence BGP return path selection in a service provider environment
- Describe the function of AS-path prepending and how you can use it to facilitate proper return path selection
- Identify design considerations when you are implementing AS-path prepending to influence return path selection
- Identify the Cisco IOS commands required to configure AS-path prepending in a multihomed network
- Identify the Cisco IOS commands required to monitor the operation of AS-path prepending
- Describe the concerns with using AS-path filters when neighboring autonomous systems require AS-path prepending

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What is AS-path prepending?

A) when a router, sending a BGP update, adds the AS number of the router from which it received the route, to the AS-path attribute

B) when a router, sending a BGP update, adds the AS number of the router to which it is sending the route, to the AS-path attribute

C) when a router, sending a BGP update, adds its AS number to the AS-path attribute multiple times.

D) when a router uses the AS-path attribute in route selection

Copyright © 2003, Cisco Systems, Inc.

Q2) AS path will be the route selection criterion used when which of the following is true?

A) It is the first criterion used in BGP route selection.

B) When there is no difference in weight, local preference, or route origination.

C) When the multi-exit discriminator is identical on the candidate routes.

D) The weight, local preference, MED, and origin attributes must be identical before the AS-path attribute is used for route selection.

Q3) What command do you use to manipulate the AS-path attribute?

A) the global configuration command, **set as-path prepend** *as-number*

B) the router configuration command, **set as-path prepend** *as-number*

C) **set as-path prepend** *as-number* in a route-map

D) the interface global command, **set as-path prepend** *as-number*

Q4) Given the following configuration from a router in AS 347, advertising network 11.0.0.0/8 to an EBGP neighbor 2.0.0.2 in AS 529:

```
route-map addAS permit 10
       set as-path prepend 347 347 347

router bgp 347
       neighbor 2.0.0.2 remote-as 529
       neighbor 2.0.0.2 route-map addAS out
```

What are the contents of the AS-path attribute for route 11.0.0.0/8 on a router residing in AS 529?

A) 347 347 347

B) 347 347 347 347

C) 529 347 347 347

D) 529 347 347 347 347

 Copyright © 2003, Cisco Systems, Inc.

Q5) Why do network administrators need to use AS-path prepending?

A) AS-path prepending allows a customer to potentially influence return path route selection.

B) AS-path prepending is used on a customer router to control outgoing route updates.

C) Service providers use AS-path prepending to control incoming updates from a customer AS.

D) AS-path prepending is used between service providers who are both connected to a customer AS, to determine who will be the primary link to the customer.

Q6) How does AS-path prepending affect a router?

A) AS-path prepending is simply a term used to describe when a router uses the AS-path attribute in route selection and hence does not affect router resources.

B) The longer the AS-path attribute attached to BGP updates, the more router memory requirements increase.

C) AS-path prepending does not impact the router because Cisco IOS software recognizes that AS-path prepending is in use and stores a single AS number with a pointer to the number of AS-path prepends.

D) AS-path prepending causes the router to operate in process-switching mode because the BGP update must be stored, manipulated, and then rewritten to accommodate for the new AS-path attribute.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

BGP Multi-Exit Discriminator

Overview

This lesson discusses how to influence Border Gateway Protocol (BGP) route selection by setting the BGP multi-exit discriminator (MED) attribute of outgoing BGP routes. The MED attribute is a hint to external neighbors about the preferred path into an autonomous system (AS) when there are multiple entry points into the AS. Two methods used to set the MED attribute are discussed in this lesson as follows: default MED and setting the MED attribute with route-maps. In addition to basic MED attribute configuration, advanced commands to manipulate MED properties are discussed. This lesson also explains how to monitor and troubleshoot the BGP table to verify correct MED configuration and properly influenced path selection.

Importance

When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. It is equally important that the return path selected be the optimum return path into the AS. The “optimum” or “best” route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides a tool for administrators to influence route selection, the MED attribute.

Objectives

Upon completing this lesson, you will be able to:

- Describe how MED can be used to facilitate proper return path selection
- Describe MED propagation inside and between autonomous systems
- Identify the Cisco IOS® commands required to configure default BGP MED on a router
- Identify the Cisco IOS commands required to configure BGP MED using route-maps

- Identify the Cisco IOS commands required to configure advanced MED properties
- Identify the Cisco IOS commands required to monitor BGP MED
- Identify the Cisco IOS commands required to troubleshoot BGP MED

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

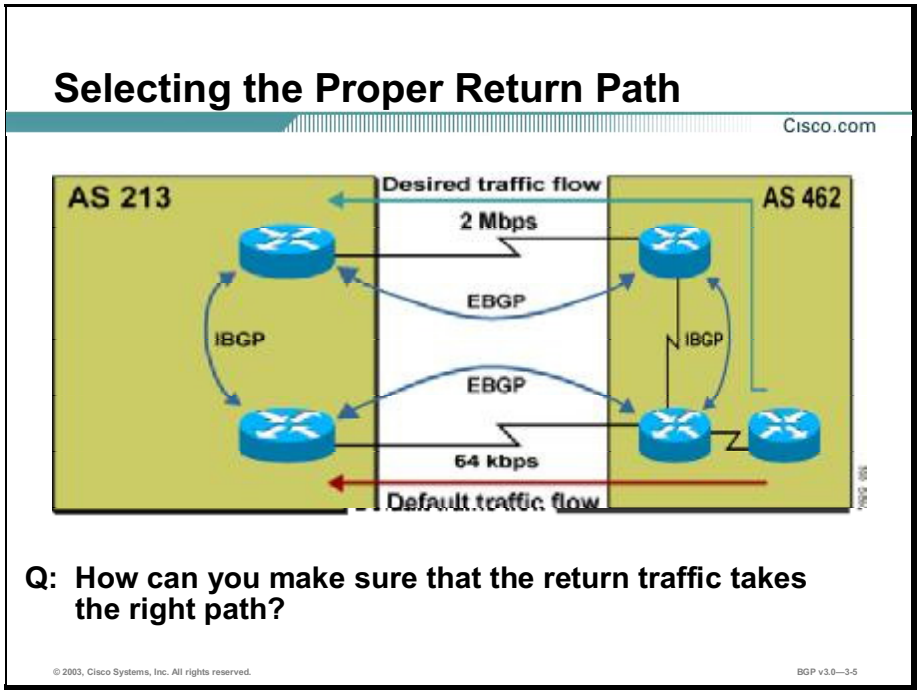
This lesson includes these topics:

- Overview
- Selecting the Proper Return Path
- MED Propagation in a BGP Network
- Changing Default MED
- Changing MED with Route-Maps
- Advanced MED Configuration
- Monitoring MED
- Troubleshooting MED
- Summary
- Assessment (Lab): BGP Multi-Exit Discriminator

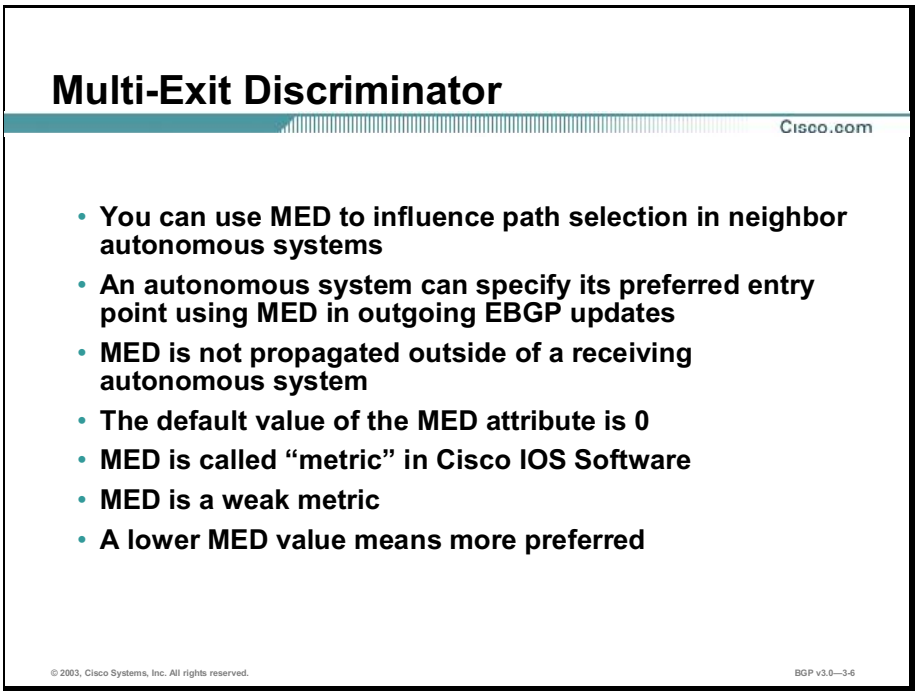
 Copyright © 2003, Cisco Systems, Inc.

Selecting the Proper Return Path

This topic describes how you can use the MED attribute to facilitate proper return path selection.



When multiple connections between providers are required, BGP attributes such as weight and local preference solve only half the problem: how to choose the right path out of the AS. This topic focuses on the second, more complex half of the same problem: how to influence neighboring autonomous systems to choose the correct return path back into the AS.


Multi-Exit Discriminator
Cisco.com
• You can use MED to influence path selection in neighbor autonomous systems
• An autonomous system can specify its preferred entry point using MED in outgoing EBGP updates
• MED is not propagated outside of a receiving autonomous system
• The default value of the MED attribute is 0
• MED is called “metric” in Cisco IOS Software
• MED is a weak metric
• A lower MED value means more preferred
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-6

You can apply the MED attribute on outgoing updates to a neighboring AS to influence the route selection process in that AS. The MED attribute is only useful when there are multiple entry points into an AS.

The MED attribute, sent to an external neighbor, will only be seen within that AS. An AS that receives a route containing the MED attribute will not advertise that MED beyond its local AS.

The default value of the MED attribute is 0. In addition, a lower value of MED is more preferred.

The MED attribute is considered a “weak” metric. In contrast with weight and local preference, a router will prefer a path with the smallest MED value but only if weight, local preference, AS path, and origin code are equal. Using MED may not yield the expected result if the neighboring AS modifies any of the stronger BGP route selection mechanisms.

Note The term used in Cisco IOS software for MED is “metric.” The term “metric” also applies to the **set** command used in route-maps as well as all **show** and **debug** commands.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What is the typical application of the MED attribute?

A) to influence path selection out of an originating AS

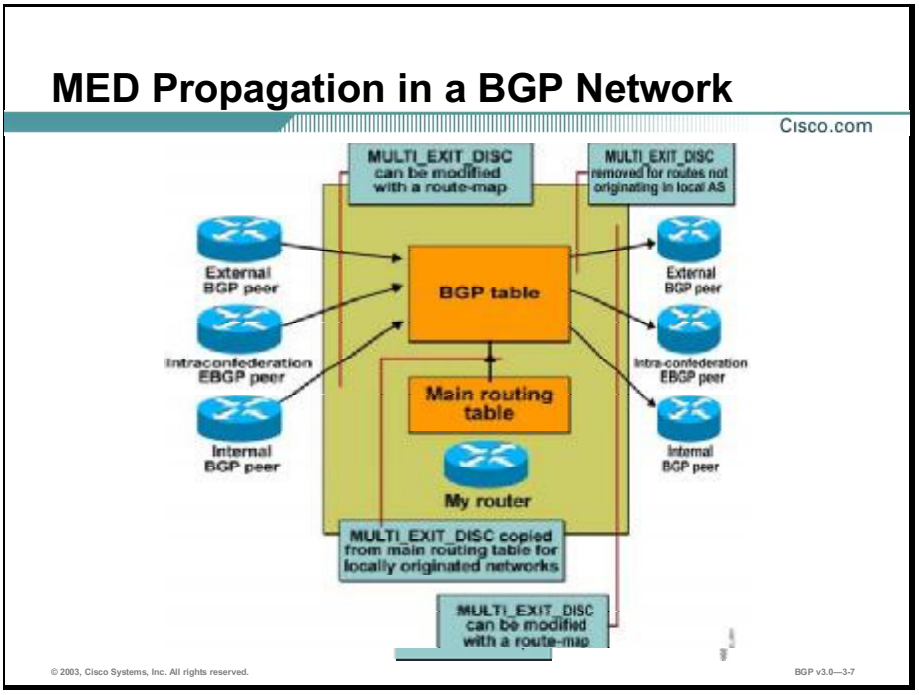
B) to provide a strong metric to select the best path when multiple routes exist

C) to have a BGP attribute traversing many autonomous systems while influencing path selection

D) to influence the return path of traffic back into an AS

MED Propagation in a BGP Network

This topic describes how the value of the MED attribute changes inside a BGP AS and between different BGP autonomous systems.


MED Propagation in a BGP Network
Cisco.com
MULTI_EXIT_DISC can be modified with a route-map
MULTI_EXIT_DISC removed for routes not originating in local AS
External BGP peer
BGP table
External BGP peer
Intraconfederation EBGP peer
Intra-confederation EBGP peer
Main routing table
Internal BGP peer
Internal BGP peer
My router
MULTI_EXIT_DISC copied from main routing table for locally originated networks
MULTI_EXIT_DISC can be modified with a route-map
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-7

The figure shows how the value of the MED attribute is assigned depending upon the routing information source. A route-map must be configured on a router to manually assign a value to the MED attribute. For those networks that are also present in the BGP table, the router assigns a default value from the metric in the routing table and copies it into the MED attribute. The MED attribute is automatically removed on external sessions if the attribute did not originate in the local AS.

Practice

Q1) What happens to the MED attribute when it is sent to EBGP peers?

A) The value of MED is ignored and reset to 0.

B) All MED values are carried into an AS but are removed by the neighbor AS.

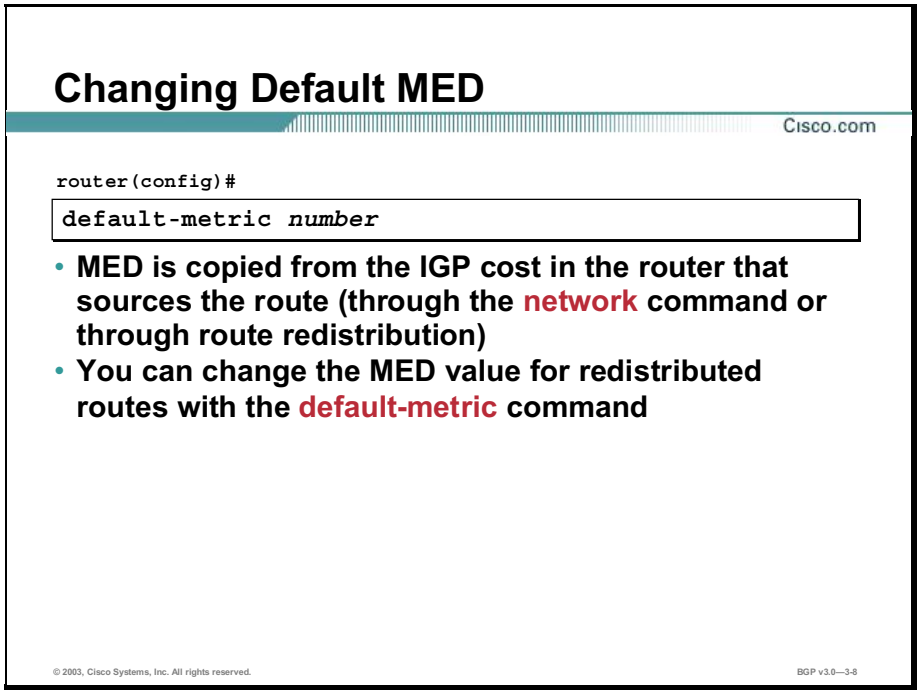
C) The MED is carried into and sent to all EBGP neighbors.

D) MED values originating in the AS are carried into the neighboring AS but do not leave it.

 Copyright © 2003, Cisco Systems, Inc.

Changing Default MED

This topic lists the Cisco IOS commands required to configure changes to the default BGP MED on a Cisco IOS router.


Changing Default MED
Cisco.com
router(config)#
default-metric number
• MED is copied from the IGP cost in the router that sources the route (through the network command or through route redistribution)
• You can change the MED value for redistributed routes with the default-metric command
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-8

MED is not a mandatory attribute, and there is no MED attribute attached to a route by default. The only exception is if the router is originating networks that have an exact match in the routing table (through the **network** command or through redistribution). In that case, the router uses the metric in the routing table as the MED attribute value.

Using the **default-metric** command in BGP configuration mode will cause all redistributed networks to have the specified MED value.

default-metric (BGP)

To set the default metric value (MED) for BGP routes, use the **default-metric** command in router configuration mode.

default-metric *number*

To return to the default state, use the **no** form of this command

no default-metric *number*

Syntax Description

number	Default metric value appropriate for the specified routing protocol

Practice

Q1) What is the default value of MED?

A) 0

B) 100

C) 32768

D) MED is unused unless explicitly defined.

Q2) What is the default value of MED when routes are redistributed from another routing protocol?

A) MED is set to 0 for all redistributed routes.

B) MED is set to 100 for all redistributed routes.

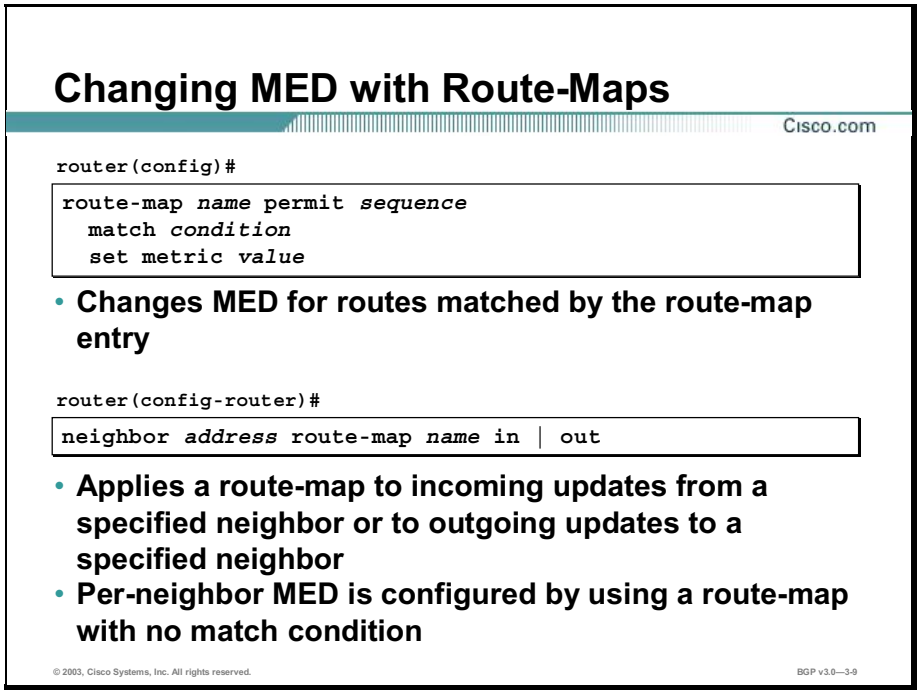
C) There is no MED attribute attached to a redistributed route.

D) MED is not attached to redistributed routes except for those matching the IP routing table.

 Copyright © 2003, Cisco Systems, Inc.

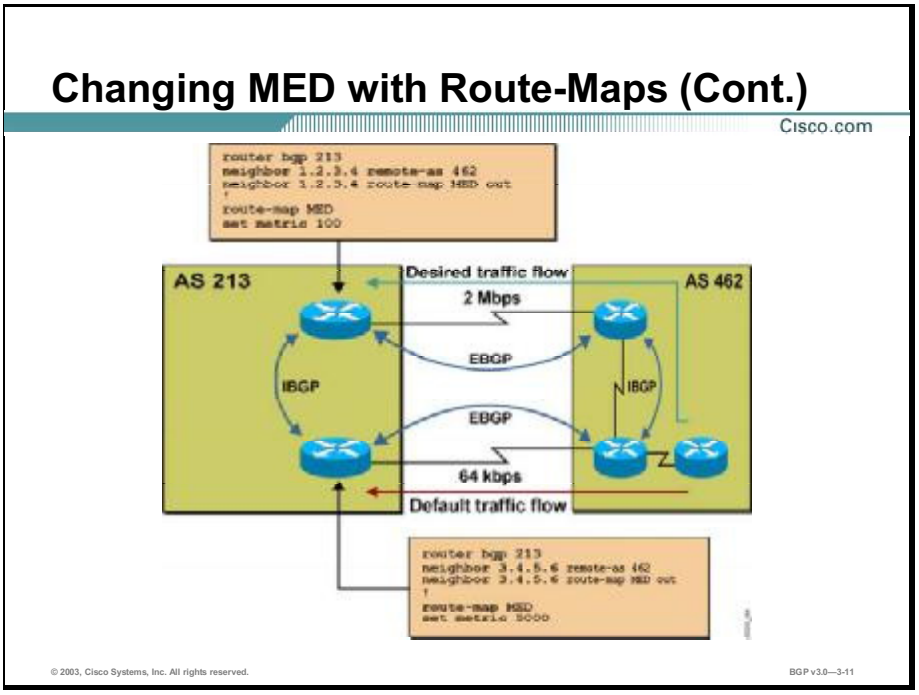
Changing MED with Route-Maps

This topic lists the Cisco IOS commands required to configure changes to the BGP MED attribute using route-map statements.


Changing MED with Route-Maps
Cisco.com
router(config)#
route-map name permit sequence
match condition
set metric value
• Changes MED for routes matched by the route-map entry
router(config-router)#
neighbor address route-map name in | out
• Applies a route-map to incoming updates from a specified neighbor or to outgoing updates to a specified neighbor
• Per-neighbor MED is configured by using a route-map with no match condition
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-9

You can use a route-map to set MED on incoming or outgoing updates. Use the **set metric** command within the route-map configuration mode to set the MED attribute.

Example



The example above shows how to set per-neighbor MED on an outgoing update. The result of this action is that the neighboring AS will prefer the upper link to AS 213. The solution, of course, relies on whether the neighboring AS is not changing the weight, local preference, AS-path, or origin code attributes in updates received from AS 213.

Practice

Q1) What are three BGP attributes that are compared before MED? (Choose three.)

A) largest weight

B) originated routes

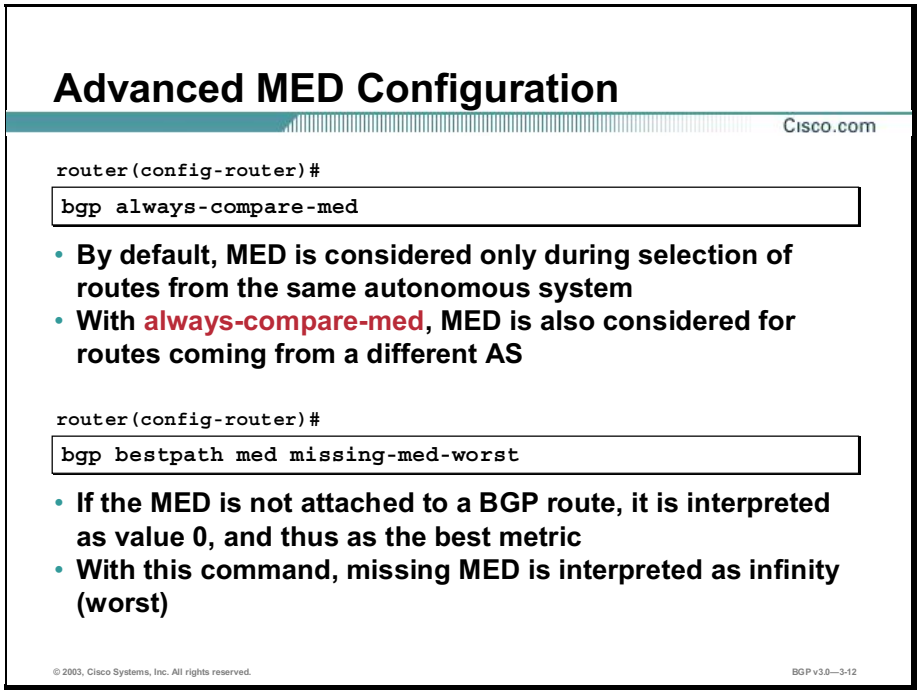
C) AS-path length

D) lowest IP address

 Copyright © 2003, Cisco Systems, Inc.

Advanced MED Configuration

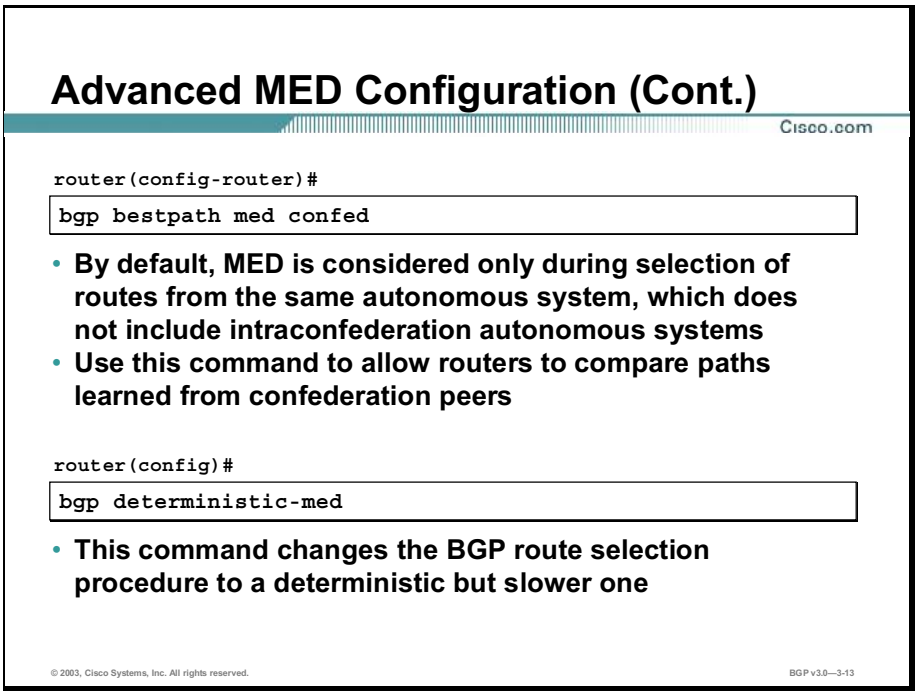
This topic lists the Cisco IOS commands required to configure advanced MED features on Cisco routers.


Advanced MED Configuration
Cisco.com
router(config-router)#
bgp always-compare-med
• By default, MED is considered only during selection of routes from the same autonomous system
• With always-compare-med, MED is also considered for routes coming from a different AS
router(config-router)#
bgp bestpath med missing-med-worst
• If the MED is not attached to a BGP route, it is interpreted as value 0, and thus as the best metric
• With this command, missing MED is interpreted as infinity (worst)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-12

Several rules exist on when and how you should use the MED attribute:

- You should use MED in the route selection process only if both (all) paths come from the same AS. Use the **bgp always-compare-med** command to force the router to compare MED even if paths come from different autonomous systems. You need to enable this option in the entire AS; otherwise routing loops can occur.

- According to a BGP standard describing MED, you should regard a missing MED attribute as an infinite value. Cisco IOS software, on the other hand, regards a missing MED attribute as having value of 0. Use the **bgp bestpath med missing-med-worst** command when combining different vendor equipment. An even better solution is to make sure that every update carries a MED attribute.


Advanced MED Configuration (Cont.)
Cisco.com
router(config-router)#
bgp bestpath med confed
• By default, MED is considered only during selection of routes from the same autonomous system, which does not include intraconfederation autonomous systems
• Use this command to allow routers to compare paths learned from confederation peers
router(config)#
bgp deterministic-med
• This command changes the BGP route selection procedure to a deterministic but slower one
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-13

You must use the command **bgp bestpath med confed** when you use MED within a confederation to influence the route selection process. A router will compare MED values for those routes that originate in the confederation.

When enabling deterministic MED comparison, a router will compare MED values before it considers BGP route type (external or internal) and Interior Gateway Protocol (IGP) metric to the next-hop address. The router will compare MED values immediately after the AS-path length.

Note	Cisco recommends enabling the **bgp deterministic-med** command in all new network rollouts. For existing networks, you must deploy the command either on all routers at the same time or incrementally, with care to avoid possible Internal Border Gateway Protocol (IBGP) routing loops.

 Copyright © 2003, Cisco Systems, Inc.

Example

The following example demonstrates how the **bgp deterministic-med** and **bgp always-compare-med** commands can influence MED-based path selection. Consider the following BGP routes for network 172.16.0.0/16 in the order they are received:

entry 1: AS(PATH) 65500, med 150, external, rid 192.168.13.1

entry 2: AS(PATH) 65100, med 200, external, rid 1.1.1.1

entry 3: AS(PATH) 65500, med 100, internal, rid 192.168.8.4

Note BGP compares multiple routes to a single destination in pairs starting with the newest entry and moving toward the oldest entry (starting at top of the list and moving down). For example, entry 1 and entry 2 are compared. The better of these two is then compared to entry 3, and so on.

In the case where both commands are disabled, BGP compares entry 1 and entry 2. Entry 2 is chosen as the best of these two because it has a lower router-ID. The MED is not checked because the paths are from a different neighbor AS. Next, entry 2 is compared to entry 3. BGP chooses Entry 2 as the best path because it is external.

In the case where **bgp deterministic-med** is disabled and **bgp always-compare-med** has been enabled, BGP compares entry 1 to entry 2. These entries are from different autonomous systems, but because the **bgp always-compare-med** command is enabled, MED is used in the comparison. Entry 1 is the better of these two entries because it has a lower MED value. Next, BGP compares entry 1 to entry 3. The MED is checked again because the entries are now from the same AS. BGP chooses entry 3 as the best path.

In the case where **bgp deterministic-med** has been enabled and **bgp always-compare-med** has been disabled, BGP groups routes from the same AS together and compares the best entries of each group. The BGP table looks like the following:

entry 1: AS(PATH) 65100, med 200, external, rid 1.1.1.1

entry 2: AS(PATH) 65500, med 100, internal, rid 192.168.8.4

entry 3: AS(PATH) 65500, med 150, external, rid 192.168.13.1

There is a group for AS 65100 and a group for AS 65500. BGP compares the best entries for each group. Entry 1 is the best of its group because it is the only route from AS 100. BGP compares entry 1 to the best of group AS 65500, entry 2 (because it has the lowest MED). Next, BGP compares entry 1 to entry 2. Because the two entries are not from the same neighbor AS, the MED is not considered in the comparison. The External Border Gateway Protocol (EBGP) route wins over the IBGP route, making entry 1 the best route.

If **bgp always-compare-med** were also enabled, BGP would have taken the MED into account for the last comparison and have selected entry 2 as the best path.

Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What effect does the **bgp deterministic-med** command have on BGP path selection?

A) It forces a MED attribute comparison between MED attributes from different neighbor autonomous systems.

B) It prohibits MED attribute comparison if routes are from a different AS.

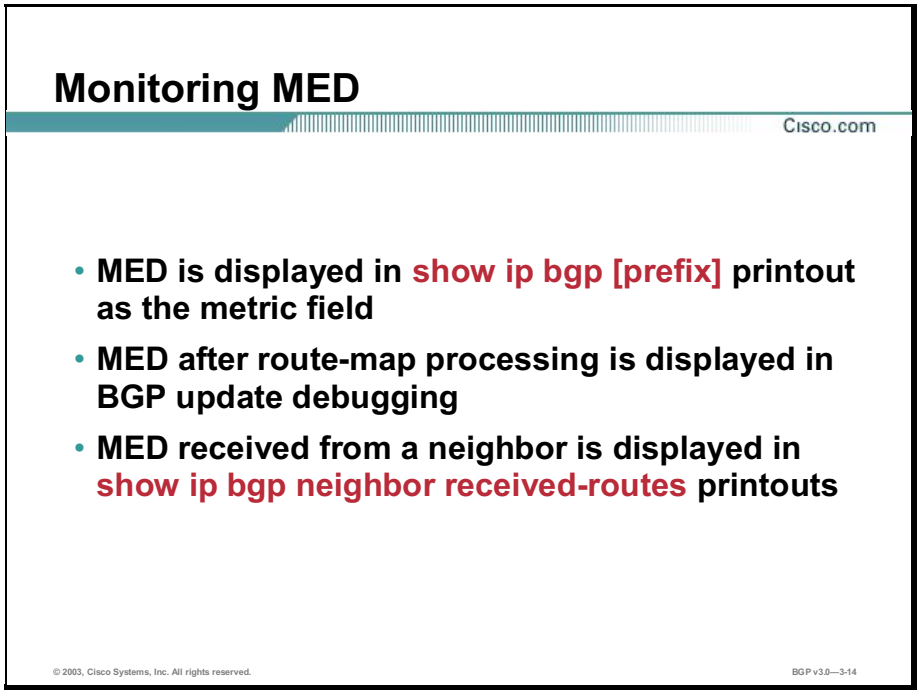
C) It ensures that an accurate MED comparison is made across all routes received from the same AS.

D) It allows the MED to determine BGP path selection above other BGP attributes.

 Copyright © 2003, Cisco Systems, Inc.

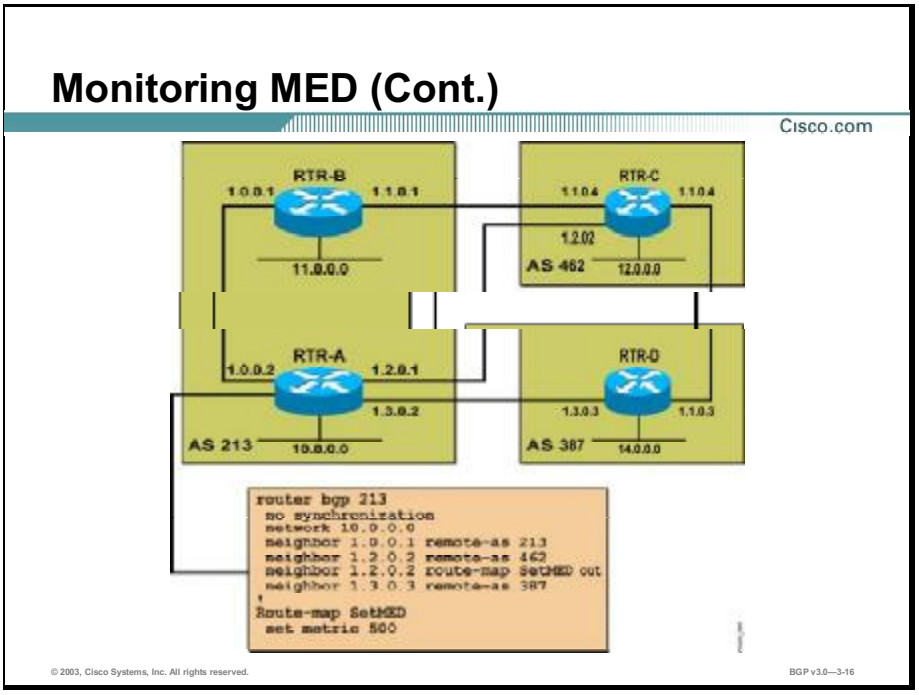
Monitoring MED

This topic lists the Cisco IOS commands required to monitor the BGP MED attribute on a Cisco router.


Monitoring MED
Cisco.com
• MED is displayed in show ip bgp [prefix] printout as the metric field
• MED after route-map processing is displayed in BGP update debugging
• MED received from a neighbor is displayed in show ip bgp neighbor received-routes printouts
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-14

All BGP-related **show** and **debug** commands display the value of the MED attribute. If the inbound soft reconfiguration feature is enabled on the router, the original MED attribute received by the router is also displayed. The following examples demonstrate command output for Cisco **show ip bgp** commands.

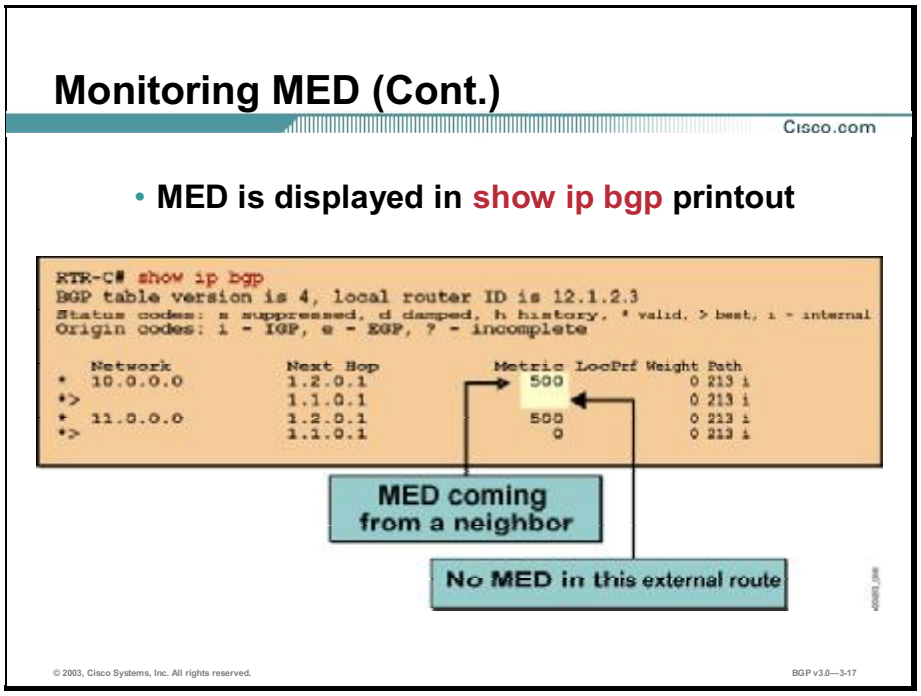
Example



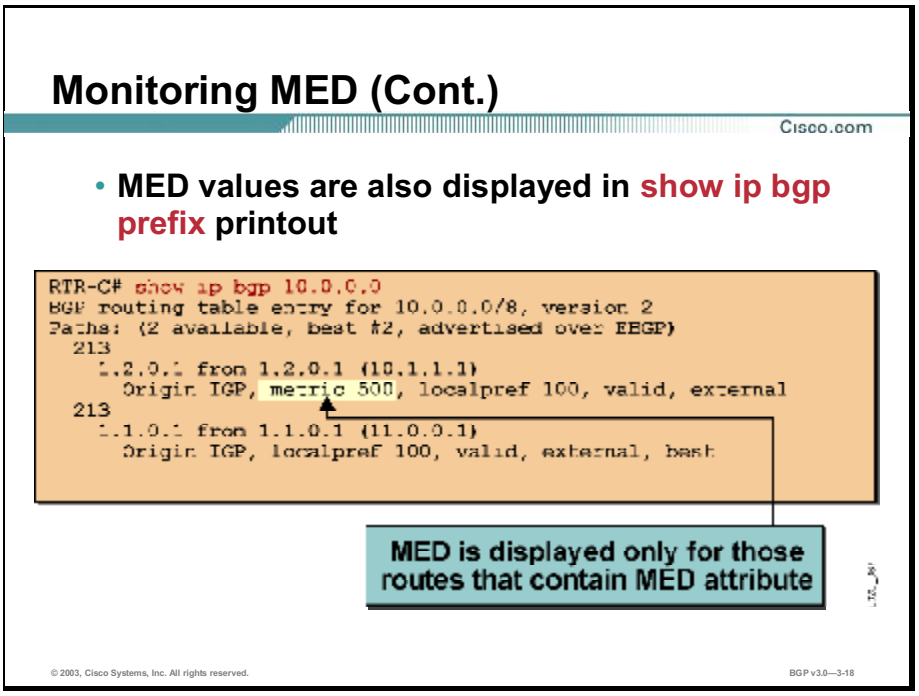
The same network as in the BGP Local Preference lesson is used in this topic to produce sample output shown on the next pages. All commands were executed on router RTR-C.

Some routing updates sent from router RTR-B are sent to router RTR-C with a MED of 500. Some updates sent from RTR-B to RTR-C have the MED set to 0, and some are without a MED attribute. Inbound soft reconfiguration is used on router RTR-C.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring MED (Cont.)
Cisco.com
• MED is displayed in show ip bgp printout
RTR-C# show ip bgp
BGP table version is 4, local router ID is 12.1.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
* 10.0.0.0 1.2.0.1 500 0 213 i
*> 1.1.0.1 0 213 i
* 11.0.0.0 1.2.0.1 500 0 213 i
*> 1.1.0.1 0 0 213 i
MED coming from a neighbor
No MED in this external route
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-17

Both networks received from router RTR-B have a MED of 500. Network 10.0.0.0/8 received from RTR-A has no MED attribute while network 11.0.0.0/8 has a MED value of 0.


Monitoring MED (Cont.)
Cisco.com
• MED values are also displayed in show ip bgp prefix printout
RTR-C# show ip bgp 10.0.0.0
BGP routing table entry for 10.0.0.0/8, version 2
Paths: (2 available, best #2, advertised over EBGP)
213
1.2.0.1 from 1.2.0.1 (10.1.1.1)
Origin IGP, metric 500, localpref 100, valid, external
213
1.1.0.1 from 1.1.0.1 (11.0.0.1)
Origin IGP, localpref 100, valid, external, best
MED is displayed only for those routes that contain MED attribute
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-18

When looking at detailed information for a specific network, you will see (**show ip bgp** *prefix*) MED only if the attribute exists.

Practice

Q1) Which two statements are true regarding MED in **show ip bgp** commands? (Choose two.)

A) MED is always displayed.

B) MED is displayed only for those routes having a MED attribute.

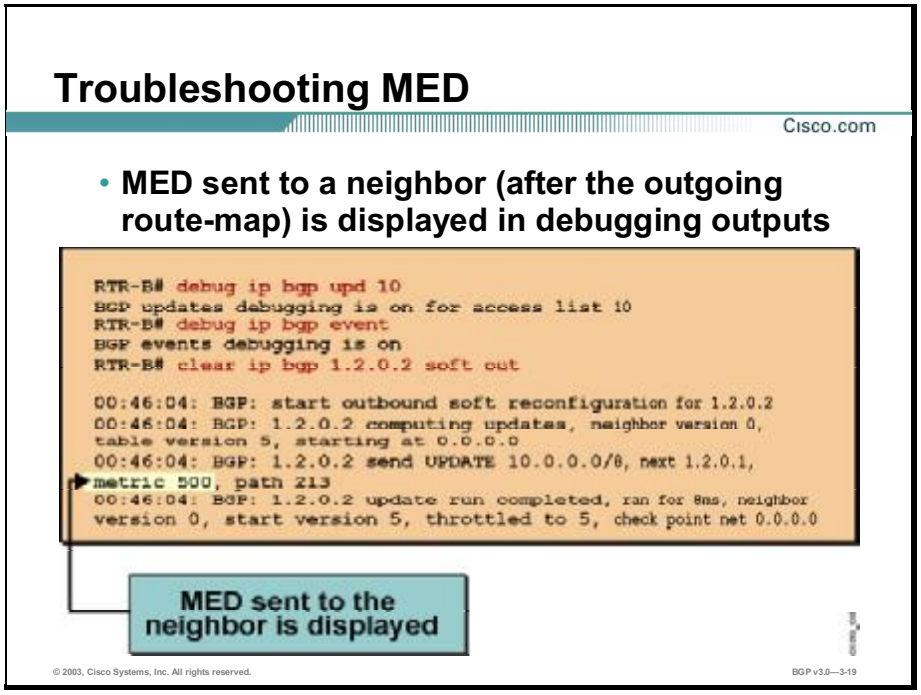
C) MED is listed as a metric value.

D) MED is listed as a MED value.

 Copyright © 2003, Cisco Systems, Inc.

Troubleshooting MED

This topic lists the Cisco IOS commands required to troubleshoot BGP MED configurations on a Cisco router.


Troubleshooting MED
Cisco.com
• MED sent to a neighbor (after the outgoing route-map) is displayed in debugging outputs
RTR-B# debug ip bgp upd 10
BGP updates debugging is on for access list 10
RTR-B# debug ip bgp event
BGP events debugging is on
RTR-B# clear ip bgp 1.2.0.2 soft out
00:46:04: BGP: start outbound soft reconfiguration for 1.2.0.2
00:46:04: BGP: 1.2.0.2 computing updates, neighbor version 0, table version 5, starting at 0.0.0.0
00:46:04: BGP: 1.2.0.2 send UPDATE 10.0.0.0/8, next 1.2.0.1, metric 500, path 213
00:46:04: BGP: 1.2.0.2 update run completed, ran for 8ms, neighbor version 0, start version 5, throttled to 5, check point net 0.0.0.0
MED sent to the neighbor is displayed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-19

In case debugging is necessary to troubleshoot a problem, MED, among other attributes, is displayed. This example shows the MED attribute set with an outgoing route-map.

Troubleshooting MED (Cont.)

Cisco.com

- **MED stored in the BGP table (after the incoming route-map processing) is displayed in debugging outputs**

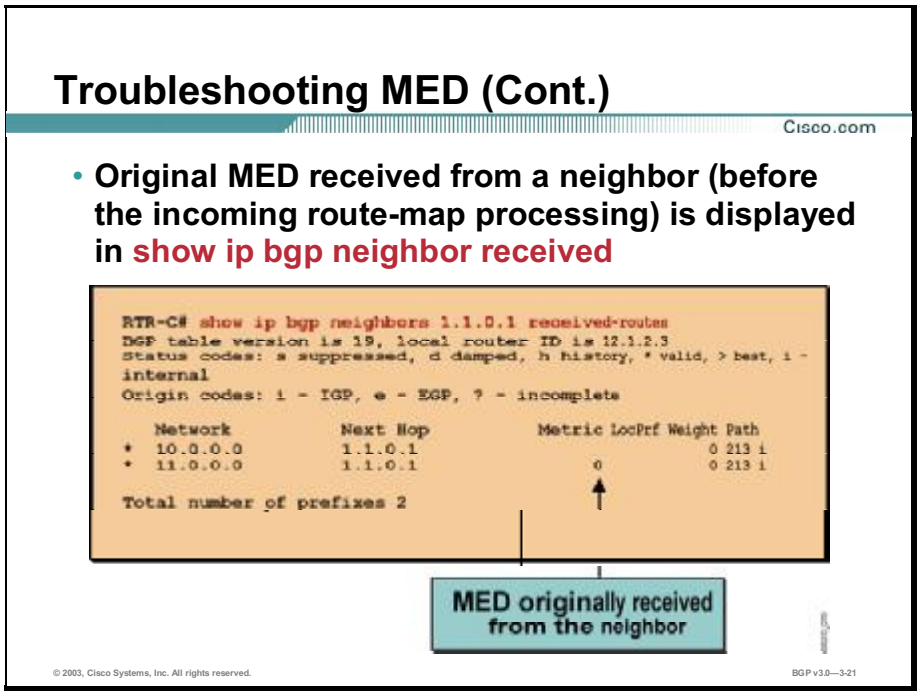
```
RTR-C# debug ip bgp update 10
BGP updates debugging is on for access list 10
RTR-C# clear ip bgp 1.2.0.1

01:03:45: BGP: 1.2.0.1 send UPDATE 10.0.0.0/8, next 1.2.0.2,
metric 0, path 462 213
01:03:45: BGP: 1.2.0.1 rcv UPDATE about 10.0.0.0/8, next hop
1.2.0.1, path 213 metric 500
```

MED stored in the BGP table is displayed

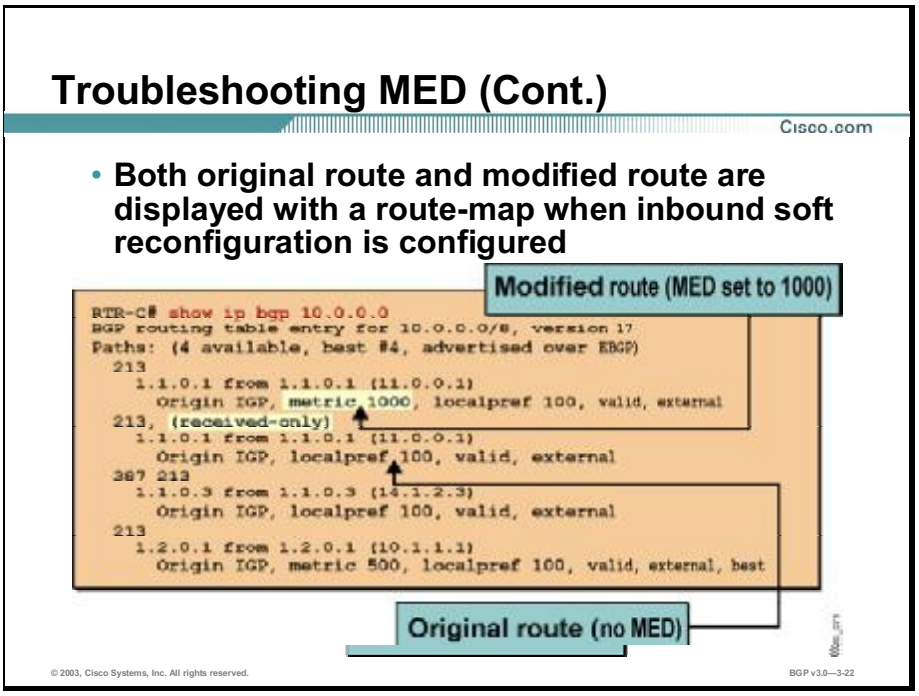
© 2003, Cisco Systems, Inc. All rights reserved.

This debugging example shows the MED attribute value after the update has been processed by an incoming route-map.

 Copyright © 2003, Cisco Systems, Inc.


Troubleshooting MED (Cont.)
Cisco.com
• Original MED received from a neighbor (before the incoming route-map processing) is displayed in show ip bgp neighbor received
RTR-C# show ip bgp neighbors 1.1.0.1 received-routes
BGP table version is 19, local router ID is 12.1.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
* 10.0.0.0 1.1.0.1 0 213 i
* 11.0.0.0 1.1.0.1 0 0 213 i
Total number of prefixes 2
MED originally received from the neighbor
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-21

To see the original MED, you need to enable soft reconfiguration on the router. The command **show ip bgp neighbor** ***address*** **received-routes** displays the original updates before any filters or route-maps have filtered or changed them.


Troubleshooting MED (Cont.)
Cisco.com
• Both original route and modified route are displayed with a route-map when inbound soft reconfiguration is configured
Modified route (MED set to 1000)
RTR-C# show ip bgp 10.0.0.0
BGP routing table entry for 10.0.0.0/8, version 17
Paths: (4 available, best #4, advertised over EBGP)
213
1.1.0.1 from 1.1.0.1 (11.0.0.1)
Origin IGP, metric 1000, localpref 100, valid, external
213, (received-only)
1.1.0.1 from 1.1.0.1 (11.0.0.1)
Origin IGP, localpref 100, valid, external
387 213
1.1.0.3 from 1.1.0.3 (14.1.2.3)
Origin IGP, localpref 100, valid, external
213
1.2.0.1 from 1.2.0.1 (10.1.1.1)
Origin IGP, metric 500, localpref 100, valid, external, best
Original route (no MED)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-22

If soft reconfiguration is enabled, the original updates to the MED attribute are available by using the **show ip bgp** *prefix* command. The original versions are marked with the **received-only** keyword and follow the version that is in the global BGP table. In the figure, the received update had no MED attribute but was later applied a value of 1000 through a route-map.

Practice

Q1) If you configure inbound soft reconfiguration with a route-map and issue the **show ip bgp** *prefix* command, which value of the MED attribute is displayed?

A) Only the original route (no MED) is displayed.

B) Both the original route and the modified route are displayed.

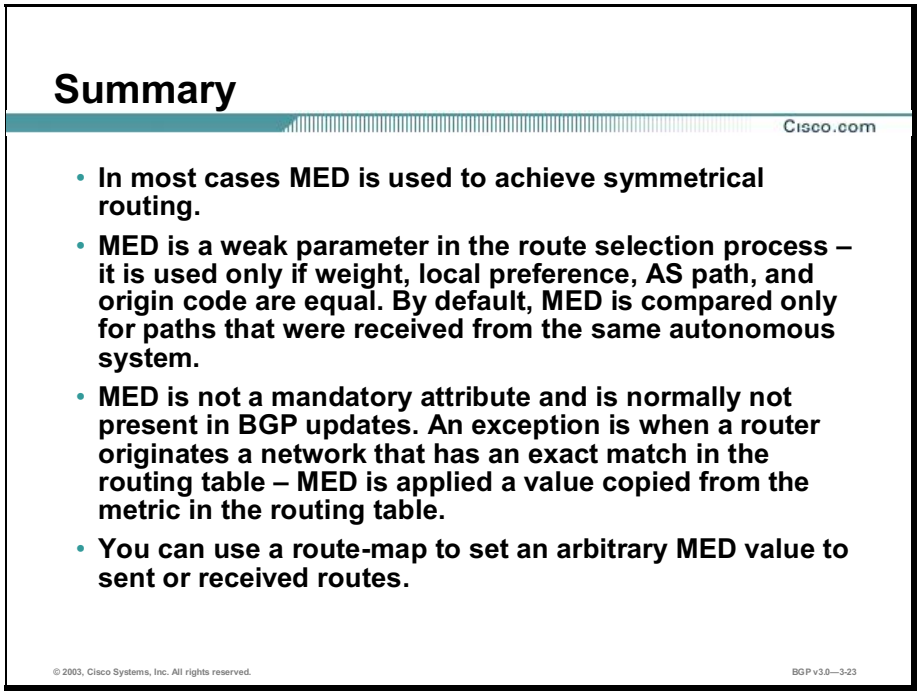
C) Only the modified route is displayed.

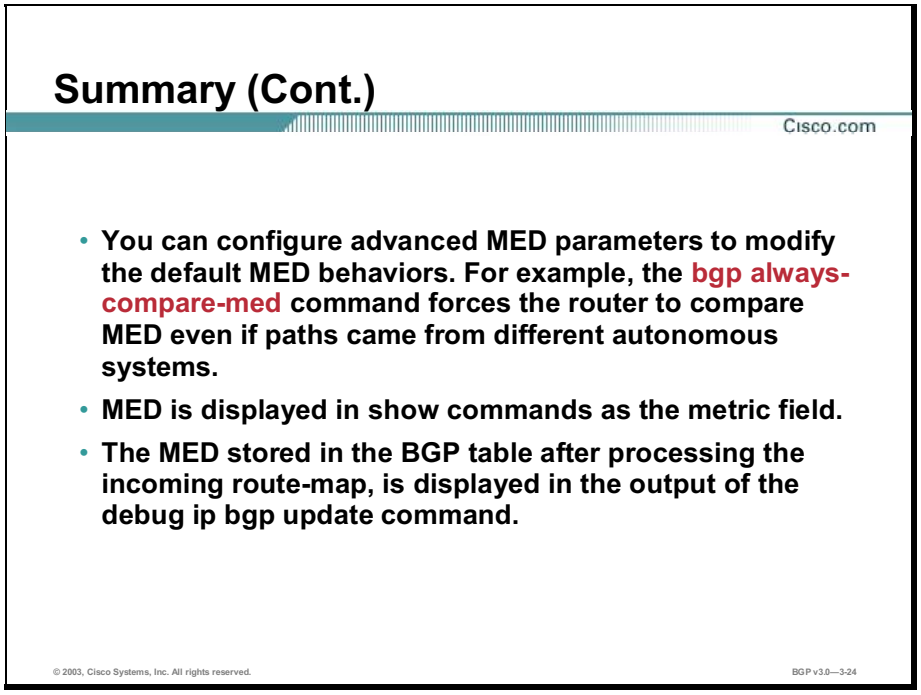
D) The MED attribute is not displayed with the **show ip bgp** *prefix* command.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• In most cases MED is used to achieve symmetrical routing.
• MED is a weak parameter in the route selection process – it is used only if weight, local preference, AS path, and origin code are equal. By default, MED is compared only for paths that were received from the same autonomous system.
• MED is not a mandatory attribute and is normally not present in BGP updates. An exception is when a router originates a network that has an exact match in the routing table – MED is applied a value copied from the metric in the routing table.
• You can use a route-map to set an arbitrary MED value to sent or received routes.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-23


Summary (Cont.)
Cisco.com
• You can configure advanced MED parameters to modify the default MED behaviors. For example, the bgp always-compare-med command forces the router to compare MED even if paths came from different autonomous systems.
• MED is displayed in show commands as the metric field.
• The MED stored in the BGP table after processing the incoming route-map, is displayed in the output of the debug ip bgp update command.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-24

Next Steps

After completing this lesson, go to:

- BGP Communities lesson

References

For additional information, refer to these resources:

- For more information on BGP MED, refer to “Border Gateway Protocol” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

- For more information on BGP communities, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/idg4/nd2003.htm#xtocid69

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: BGP Multi-Exit Discriminator

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using the MED attribute in a situation where you must support multiple connections to an Internet service provider (ISP).

After completing this exercise, you will be able to:

- Configure BGP MED using route-maps
- Monitor BGP MED

Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in | out**}	Use this command to apply a route-map to incoming or outgoing routing updates.
route-map *name* {**permit | deny**} **seq**	Define or modify an existing entry in a route-map.
set metric *metric*	Set MED in a route-map.
clear ip bgp *	Reset BGP peer.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- You have noticed that the traffic from router "Good" toward your AS sometimes passes through the WGxR1, resulting in asymmetrical routing. You will use MED to indicate to the "Good" router which exit point it should use.

Note This exercise is a continuation of the BGP Local Preference lab exercise.

- The implementation should ensure symmetrical routing. You should use backup peering only when the primary link is down. Use MED to influence the neighboring AS 20 to choose the preferred return path.

- Figure 1 shows the physical connectivity, BGP sessions, and traffic flow in the network.

 Copyright © 2003, Cisco Systems, Inc.

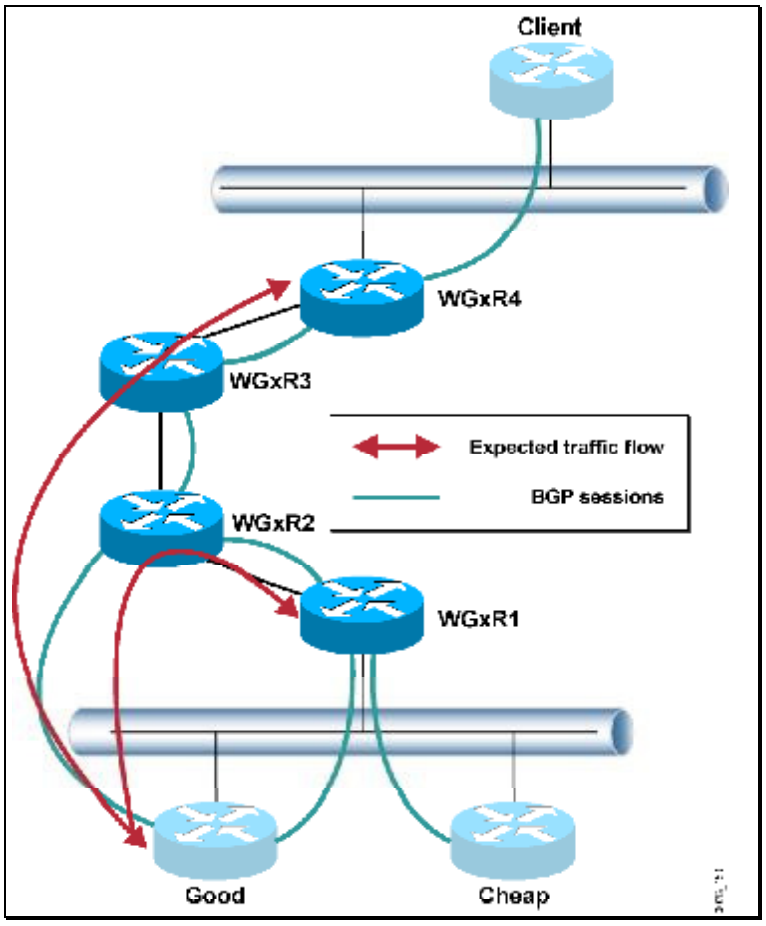


Figure 1: BGP MED in a service provider network

Exercise Procedure

Complete these steps:

Step 1 Before starting the exercise, check the BGP table on router "Good" to see if MED is present in BGP updates received from WGxR1 and WGxR2.

Configuring WGxR1 to set MED:

Step 2 Create a new route-map, which sets the BGP MED to a high value (for example, 200). Apply the route-map to outgoing updates toward the router "Good."

Configuring WGxR2 to set MED:

Step 3 Create a new route-map, which sets the BGP MED to a low value (for example, 100). Apply the route-map to outgoing updates toward the router "Good."

Copyright © 2003, Cisco Systems, Inc.

Exercise Verification

You have completed this exercise when you attain these results:

- Check the BGP table on router “Good” to see if MED is present in BGP updates received from WGxR1 and WGxR2.

```
Good#sh ip bgp
BGP table version is 46, local router ID is 199.199.199.199
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.0.0.0         192.168.20.233           0         32768 i
*  128.20.0.0       192.168.20.22            0             0 22 i
*>                  0.0.0.0                  0         32768 i
*> 128.20.12.0/24   0.0.0.0                  0         32768 i
*  128.22.0.0       192.168.20.22            0             0 22 i
*>                  0.0.0.0                  0         32768 i
*> 128.22.12.0/24   192.168.20.22            0             0 22 i
*  128.26.0.0       192.168.20.22            0             0 22 26 i
*>                  0.0.0.0                  0         32768 i
*  128.37.0.0       192.168.20.22            0             0 22 26 42
37 i
*>                  0.0.0.0                  0         32768 i
*  128.42.0.0       192.168.20.22            0             0 22 26 42
i
*>                  0.0.0.0                  0         32768 i
*  128.51.0.0       192.168.20.22            0             0 22 26 51
i
*>                  0.0.0.0                  0         32768 i
*> 128.213.0.0      0.0.0.0                  0         32768 i
*> 128.214.0.0      192.168.20.22            0             0 22 214 i
*  192.20.11.0      192.168.20.22            0             0 22 i
*>                  0.0.0.0                  0         32768 i
*> 192.20.12.0/30   0.0.0.0                  0         32768 i
*  192.22.11.0      192.168.20.22            0             0 22 i
*>                  0.0.0.0                  0         32768 i
*> 192.22.12.0/30   192.168.20.22            0             0 22 i
*  192.26.11.0      192.168.20.22            0             0 22 26 i
*>                  0.0.0.0                  0         32768 i
*  192.37.11.0      192.168.20.22            0             0 22 26 42
37 i
*>                  0.0.0.0                  0         32768 i
*  192.42.11.0      192.168.20.22            0             0 22 26 42
i
*>                  0.0.0.0                  0         32768 i
*  192.51.11.0      192.168.20.22            0             0 22 26 51
i
*>                  0.0.0.0                  0         32768 i
*  192.168.1.0      192.168.20.1                           0 22 1 i
*                   192.168.20.1           150             0 1 i
*>                  192.168.31.1            50             0 1 i
*  192.168.31.0     192.168.20.1                           0 22 1 i
*>                  192.168.31.1            50             0 1 i
*                   192.168.20.1           150             0 1 i
*> 192.213.11.0     0.0.0.0                  0         32768 i
*> 192.214.11.0     192.168.20.22            0             0 22 214 i
*  197.1.0.0/21     192.168.20.1                           0 22 1 i
*>                  192.168.31.1            50             0 1 i
*                   192.168.20.1           150             0 1 i
*  197.1.0.0/16     192.168.20.1                           0 22 1 i
*>                  192.168.31.1            50             0 1 i
*                   192.168.20.1           150             0 1 i
```

 Copyright © 2003, Cisco Systems, Inc.

```
*  197.1.8.0/22     192.168.20.1                           0 22 1 i
*>                  192.168.31.1            50             0 1 i
*                   192.168.20.1           150             0 1 i
*> 200.20.0.0/16    0.0.0.0                  0         32768 i
*> 200.22.0.0/16    192.168.20.22            0             0 22 I
```

- Use traceroute from routers WGxR4 and WGxR1 to 192.20.11.1.

```
wg1r4#traceroute 192.20.11.1
Type escape sequence to abort.
Tracing the route to 192.20.11.1

  1 192.168.1.9 32 msec 24 msec 20 msec
  2 192.168.1.5 28 msec 44 msec 41 msec
  3 192.168.31.2 60 msec 64 msec 56 msec


wg1r1#traceroute 192.20.11.1
Type escape sequence to abort.
Tracing the route to 192.20.11.1

  1 192.168.1.2 40 msec 28 msec 20 msec
  2 192.168.31.2 32 msec *  36 msec
```

- Use traceroute from router “Good” to interface loopback0 on routers WGxR4 and WGxR1.

```
Good#traceroute 197.1.7.1
Type escape sequence to abort.
Tracing the route to 197.1.7.1

  1 192.168.31.1 [AS 1] 36 msec 28 msec 24 msec
  2 192.168.1.6 [AS 1] 32 msec 36 msec 36 msec
  3 192.168.1.10 [AS 1] 52 msec *  48 msec


Good#traceroute 197.1.8.1
Type escape sequence to abort.
Tracing the route to 197.1.8.1

  1 192.168.31.1 [AS 1] 36 msec 28 msec 24 msec
  2 192.168.1.1 [AS 1] 40 msec *  36 msec
```

- Compare the two outputs of traceroute to determine if routing is symmetrical. Both trace commands should show router WGxR2 in the path.

Answer these questions:

Q1) Which parameters and attributes have to be equal before MED is compared to select the best path?

Q2) What is the default value of MED?

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

BGP Communities

Overview

This lesson discusses how to influence Border Gateway Protocol (BGP) route selection by setting the BGP community attribute on outgoing BGP routes. The community attribute provides a method to facilitate and simplify the control of routing information by grouping destinations into a specific community. Routing decisions can then be influenced based on the identity of a group, simplifying the configuration requirements of BGP when applying administrative policies.

In this lesson, BGP communities and their use to facilitate proper return path selection is discussed. The configuration details of BGP communities, and the use of community-lists and route-maps to influence route selection are also discussed. This lesson concludes by explaining how to monitor BGP community attributes.

Importance

The community attribute is a transitive optional BGP attribute, designed to group destinations and allow the easy application of administrative policies. BGP communities provide a mechanism to reduce BGP configuration complexity on a router controlling the distribution of routing information. When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. It is equally important that the return path selected be the optimum return path into the autonomous system (AS). The “optimum” or “best” route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides tools for administrators to influence route selection. The BGP community attribute is one such tool.

Objectives

Upon completing this lesson, you will be able to:

- Describe the issues of return path selection for multihomed customers
- Describe the basic qualities of BGP communities
- Describe how you can use BGP communities to facilitate proper return path selection
- List the steps required to successfully deploy communities in a BGP-based network
- Identify the Cisco IOS® commands required to configure route tagging using BGP communities
- Identify the Cisco IOS commands required to configure BGP community propagation
- Identify the Cisco IOS commands required to match routes based on attached BGP communities using community-lists
- Identify the Cisco IOS commands required to match routes based on attached BGP communities using route-maps
- Identify the Cisco IOS commands required to monitor BGP communities

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

 Copyright © 2003, Cisco Systems, Inc.

Outline

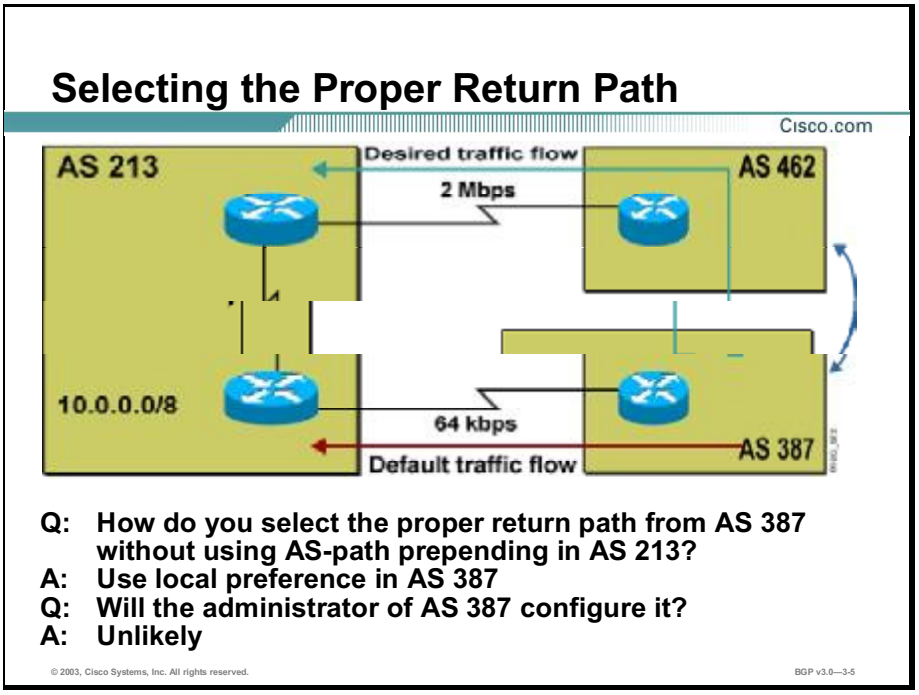
This lesson includes these topics:

- Overview
- Selecting the Proper Return Path
- BGP Communities Overview
- Using Communities
- Configuring BGP Communities
- Configuring Route Tagging with BGP Communities
- Configuring Community Propagation
- Defining BGP Community-Lists
- Matching BGP Communities in Route-Maps
- Monitoring Communities
- Summary
- Assessment (Lab): BGP Communities

Copyright © 2003, Cisco Systems, Inc.

Selecting the Proper Return Path

This topic describes the issues of return path selection for multihomed customers and why you cannot use the BGP attributes of weight, local preference, and MED to solve these issues.


Selecting the Proper Return Path
Cisco.com
AS 213
Desired traffic flow
2 Mbps
AS 462
10.0.0.0/8
64 kbps
AS 387
Default traffic flow
Q: How do you select the proper return path from AS 387 without using AS-path prepending in AS 213?
A: Use local preference in AS 387
Q: Will the administrator of AS 387 configure it?
A: Unlikely
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-5

In this example, the customer and the backup service provider would like to avoid AS-path prepending and rely on other BGP tools to properly route the return traffic over the highest-speed WAN link.

Using MED to influence the preferred return path is not possible because MED cannot be propagated across several autonomous systems. AS 387 would, therefore, receive networks from AS 213 directly with the MED attribute, but without a MED attribute from AS 462. In any case, BGP route selection would be based on the length of the AS path, and even if MED was present and used the shortest path, it would still be through the slow 64-kbps link.

The only option for solving this issue is to use local preference in AS 387. The problem with this solution is that service providers normally do not rush to implement every wish that their customers might have.

This lesson describes a solution to this case study using the transitive optional attribute called BGP community in conjunction with local preference.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What are two reasons why it is not feasible to use MED to influence return path selection when multiple autonomous systems are involved? (Choose two.)

A) because the MED attribute is designed to influence outbound path selection only

B) because the AS-path attribute would be used for path selection regardless of any configured MED value

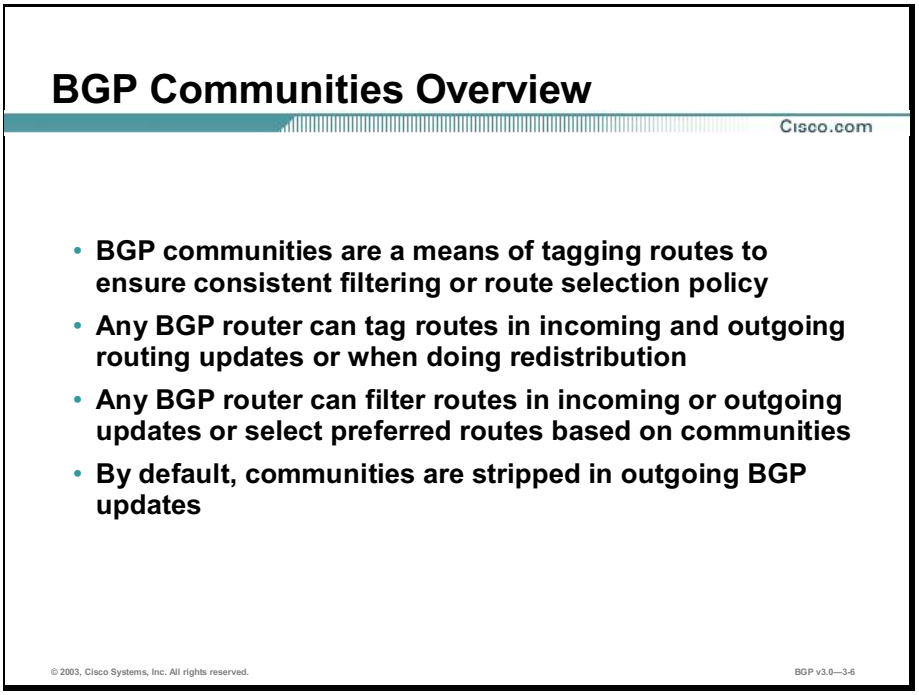
C) because the weight attribute will always be used, given that it is first in the BGP route selection process

D) because MED cannot be propagated across several autonomous systems

Copyright © 2003, Cisco Systems, Inc.

BGP Communities Overview

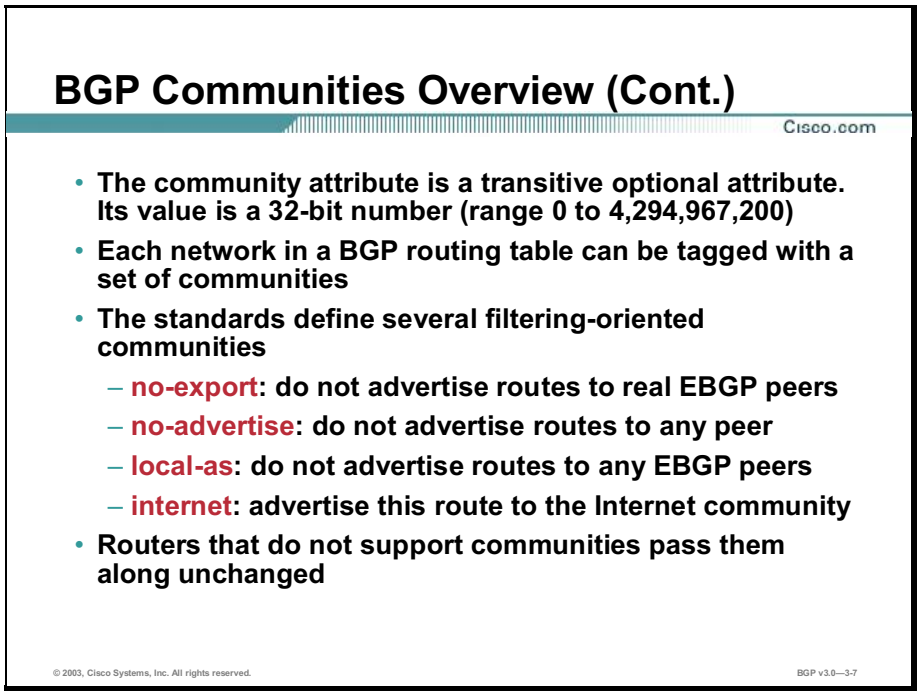
This topic describes the basic properties and fundamental qualities of BGP communities.


BGP Communities Overview
Cisco.com
• BGP communities are a means of tagging routes to ensure consistent filtering or route selection policy
• Any BGP router can tag routes in incoming and outgoing routing updates or when doing redistribution
• Any BGP router can filter routes in incoming or outgoing updates or select preferred routes based on communities
• By default, communities are stripped in outgoing BGP updates
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-6

A community is an attribute used to tag BGP routes. A router can apply it to any BGP route by using a route-map. Other routers can then perform any action based on the tag (community) that is attached to the route.

There can be more than one BGP community attached to a single route, but the routers, by default, remove communities in outgoing BGP updates.

 Copyright © 2003, Cisco Systems, Inc.


BGP Communities Overview (Cont.)
Cisco.com
• The community attribute is a transitive optional attribute. Its value is a 32-bit number (range 0 to 4,294,967,200)
• Each network in a BGP routing table can be tagged with a set of communities
• The standards define several filtering-oriented communities
– no-export: do not advertise routes to real EBGP peers
– no-advertise: do not advertise routes to any peer
– local-as: do not advertise routes to any EBGP peers
– internet: advertise this route to the Internet community
• Routers that do not support communities pass them along unchanged
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-7

The community attribute is a 32-bit transitive optional attribute that provides a way to group destinations and apply routing decisions (accept, prefer, redistribute, and so on) according to communities. A set of community values has been pre-defined. When a router receives a route marked with a pre-defined community, it will perform a specific, pre-defined action based on that community setting as follows:

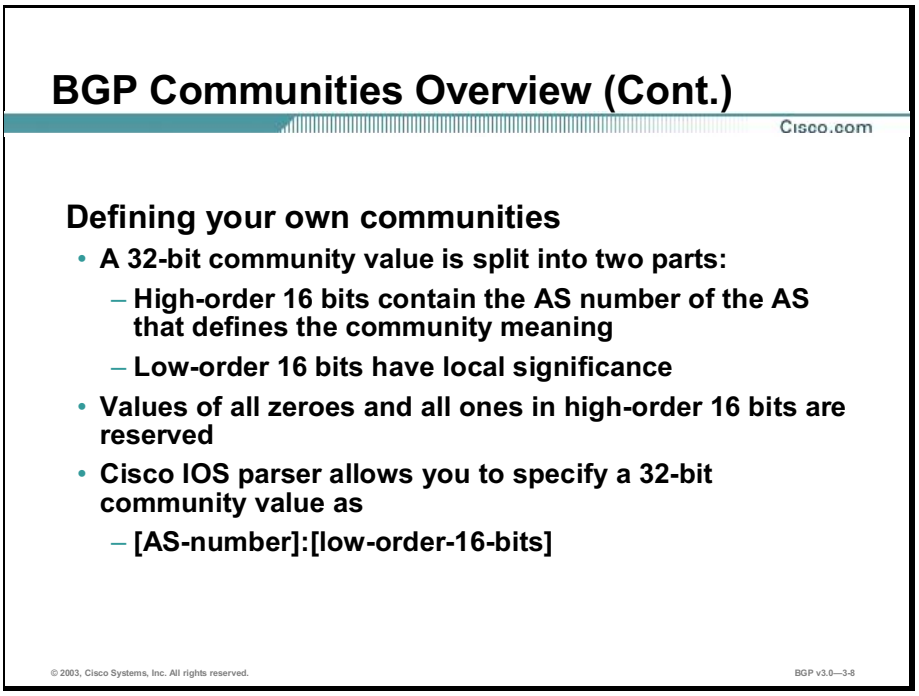
- **no-export:** If a router receives an update carrying this community, it will not propagate it to any external neighbors except intraconfederation external neighbors. This is the most widely used predefined community attribute.

- **no-advertise:** If a router receives an update carrying this community, it will not forward it to any neighbor.

- **local-as:** This community has a similar meaning to **no-export**, but it keeps a route within the local AS (or member-AS within the confederation). The route is not sent to external BGP neighbors or to intraconfederation external neighbors.

- **internet:** Advertise this route to the Internet community. All routers belong to it.

Routers that do not support the community attribute will pass the attribute to other neighbors because it is a transitive attribute.


BGP Communities Overview (Cont.)
Cisco.com
Defining your own communities
• A 32-bit community value is split into two parts:
– High-order 16 bits contain the AS number of the AS that defines the community meaning
– Low-order 16 bits have local significance
• Values of all zeroes and all ones in high-order 16 bits are reserved
• Cisco IOS parser allows you to specify a 32-bit community value as
– [AS-number]:[low-order-16-bits]
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-8

Community attributes are usually used between neighboring autonomous systems. To make sure that the BGP communities are globally unique, a public AS number should be part of the community value. For this reason, you can enter the community value as two 16-bit numbers separated by a colon. The first number (high-order 16 bits) should be the AS number of the AS defining the community value, and the second number is a value that is assigned a certain meaning (i.e., translation of a community value into local preference in the neighboring AS).

Communities can also be used internally within an AS (to ensure AS-wide routing policy) in which case the first 16 bits should contain the AS number of the local AS.

Practice

Q2) What is the purpose of the community attribute?

A) to filter incoming BGP route updates

B) to update the BGP table with incoming BGP routes

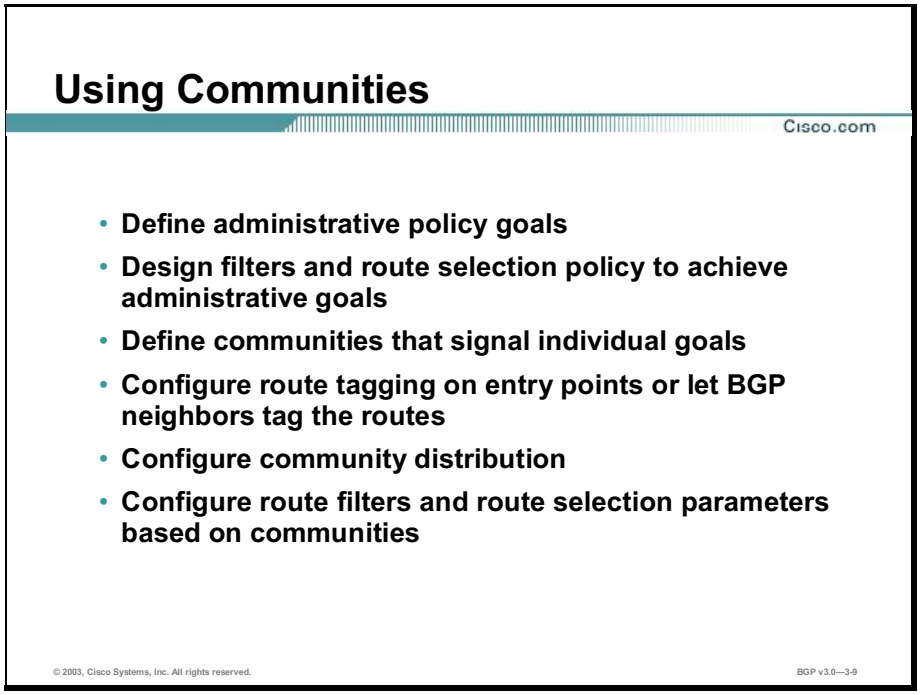
C) to facilitate selection of the optimum AS exit path

D) to group destinations by tagging BGP updates

 Copyright © 2003, Cisco Systems, Inc.

Using Communities

This topic describes how BGP communities can facilitate proper return path selection.


Using Communities
Cisco.com
• Define administrative policy goals
• Design filters and route selection policy to achieve administrative goals
• Define communities that signal individual goals
• Configure route tagging on entry points or let BGP neighbors tag the routes
• Configure community distribution
• Configure route filters and route selection parameters based on communities
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-9

Designing a BGP solution around BGP communities usually requires the following steps:

Step 1 Define administrative policy goals that you need to implement.

Step 2 Define the filters and route selection policy that will achieve the required goal(s).

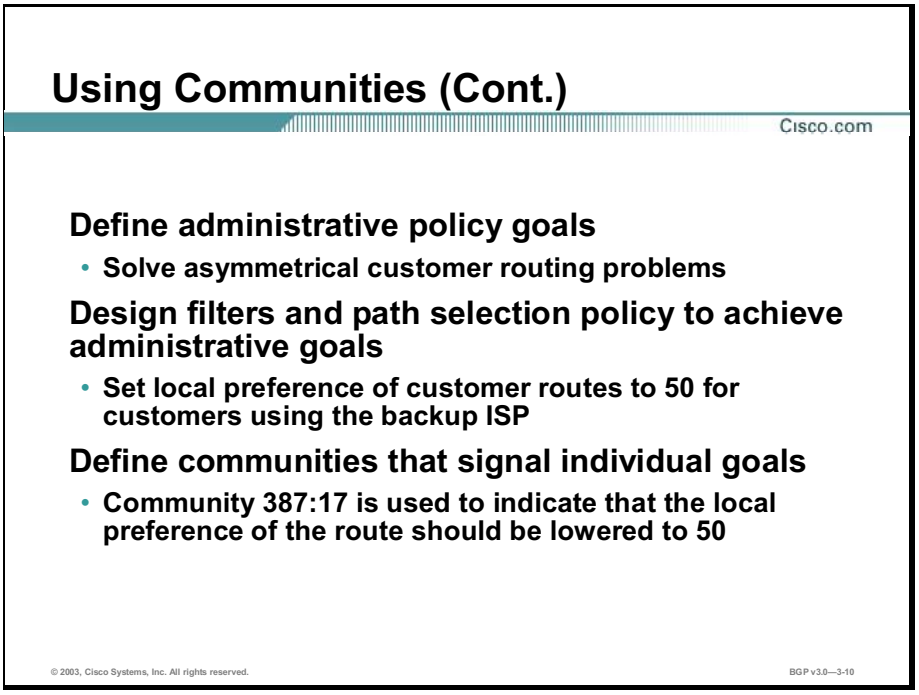
Step 3 Assign a community value to each goal.

Step 4 Apply communities on incoming updates from neighboring autonomous systems or tell the neighbors to set the communities themselves.

Step 5 Enable community distribution throughout your AS to allow community propagation.

Step 6 Match communities with route-maps and route filters, change BGP attributes, or influence route selection process based on the communities attached to the BGP routes.

Example


Using Communities (Cont.)
Cisco.com
Define administrative policy goals
• Solve asymmetrical customer routing problems
Design filters and path selection policy to achieve administrative goals
• Set local preference of customer routes to 50 for customers using the backup ISP
Define communities that signal individual goals
• Community 387:17 is used to indicate that the local preference of the route should be lowered to 50
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-10

The following table is an example of how you can define goals and assign communities to them:

Goal	Community value
Set local preference of 50	387:17
Set local preference of 150	387:18
Prepend AS path once when sending the network to external neighbors	387:21
Prepend AS path twice when sending the network to external neighbors	387:22
Prepend AS path three times when sending the network to external neighbors	387:23

All customers of the service provider should know this list so that they can use the BGP communities without having to discuss their use with the service provider.

Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Does the community attribute have any influence on BGP path selection?

A) No, communities are simply tags that are applied to BGP routes.

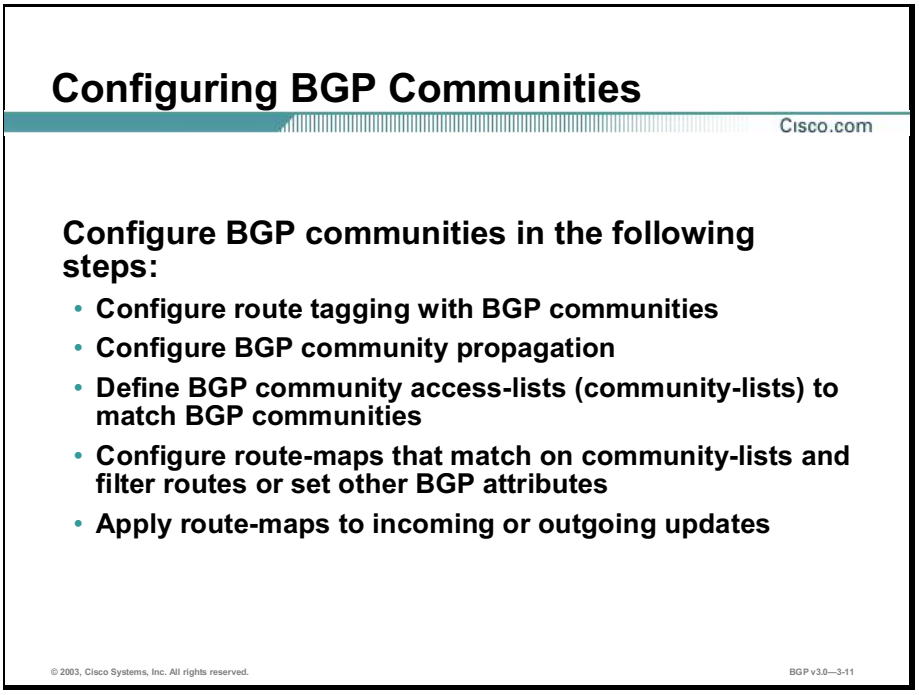
B) No, communities are nontransitive attributes.

C) Yes, BGP paths are selected based on the value in the community tag.

D) Yes, the community attribute is part of the BGP route selection process.

Configuring BGP Communities

This topic lists the steps required to successfully deploy BGP communities in a BGP-based network.


Configuring BGP Communities
Cisco.com
Configure BGP communities in the following steps:
• Configure route tagging with BGP communities
• Configure BGP community propagation
• Define BGP community access-lists (community-lists) to match BGP communities
• Configure route-maps that match on community-lists and filter routes or set other BGP attributes
• Apply route-maps to incoming or outgoing updates
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-11

Configuration steps when you are using communities include:

- Setting communities, which requires a route-map.
- Enabling community propagation per neighbor for all internal neighbors. If communities are sent to external neighbors, you must enable community propagation for external neighbors.
- Creating community-lists to be used within route-maps to match on community values.
- Creating route-maps where community-lists are used to match on community values. You can then use route-maps to filter based on community values or to set other parameters or attributes (for example, local preference, MED, or AS-path prepending).
- Applying route-maps to incoming or outgoing updates.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What Cisco IOS feature must you use to set the BGP community attribute?

A) distribute-list

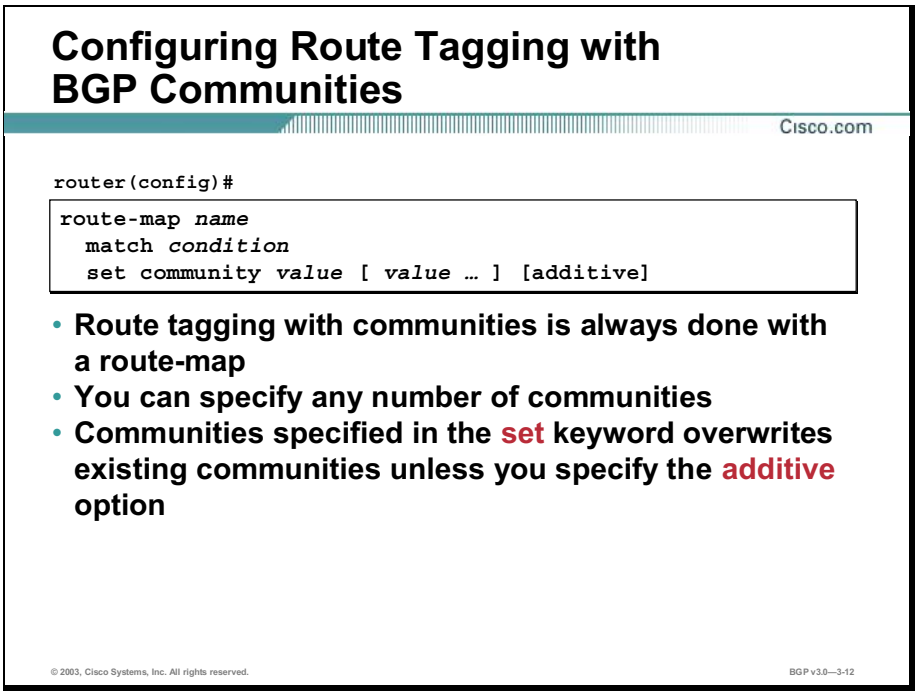
B) filter-list

C) route-map

D) access control list

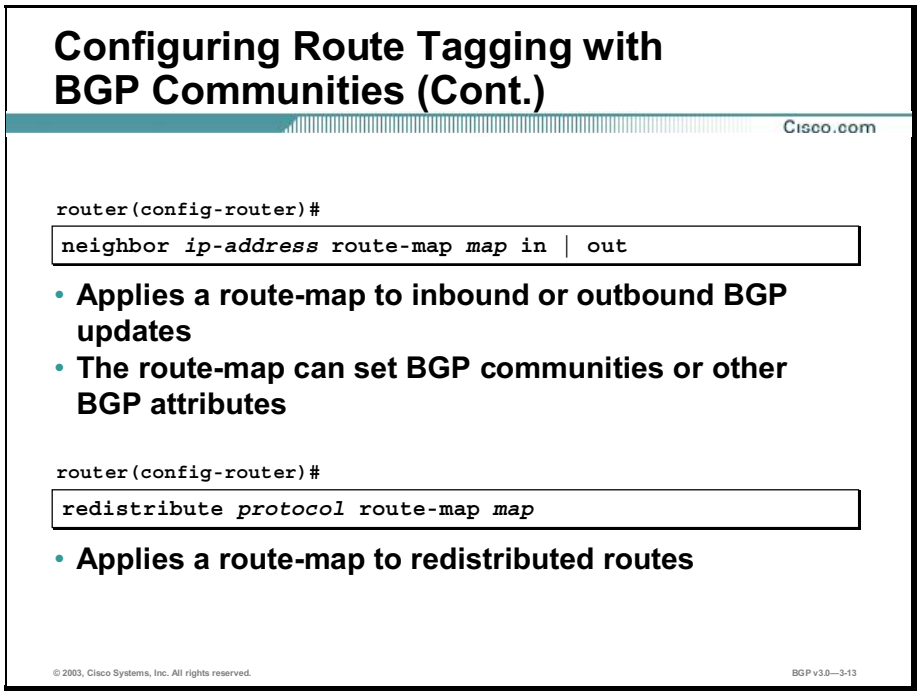
Configuring Route Tagging with BGP Communities

This topic lists the Cisco IOS commands required to configure route tagging using BGP communities.


Configuring Route Tagging with BGP Communities
Cisco.com
router(config)#
route-map name
match condition
set community value [value …] [additive]
• Route tagging with communities is always done with a route-map
• You can specify any number of communities
• Communities specified in the set keyword overwrites existing communities unless you specify the additive option
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-12

In a route-map configuration mode, you should use the **set community** command to attach a community attribute (or a set of communities) to a route. You can attach up to 32 communities to a single route with one route-map **set** statement. If the keyword **additive** is used, the original communities are preserved and the router simply appends the new communities to the route. Omitting the **additive** keyword results in overwriting any original community attributes.

 Copyright © 2003, Cisco Systems, Inc.


Configuring Route Tagging with BGP Communities (Cont.)
Cisco.com
router(config-router)#
neighbor ip-address route-map map in | out
• Applies a route-map to inbound or outbound BGP updates
• The route-map can set BGP communities or other BGP attributes
router(config-router)#
redistribute protocol route-map map
• Applies a route-map to redistributed routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-13

You can apply a route-map to incoming or outgoing updates. You can also use it with redistribution from another routing protocol.

Note	A route-map is a filtering mechanism that has an “implicit deny” for all networks not matched in any route-map statement. If a route-map is not intended to filter routes, then you should add another route-map statement at the end to permit all remaining networks without changing it (no **match** and no **set** commands are used within that route-map statement).

Originally, Cisco IOS software accepted and displayed BGP community values as a single 32-bit value in a digital format. Newer Cisco IOS versions support the new format, where you can set or view a community as two colon-separated 16-bit numbers.

The global command **ip bgp-community new-format** is recommended on all routers whenever communities contain the AS number.

After being converted, configuration files with communities in the new as:nn format are not compatible with older versions of Cisco IOS software. For example:

router# show ip bgp 6.0.0.0

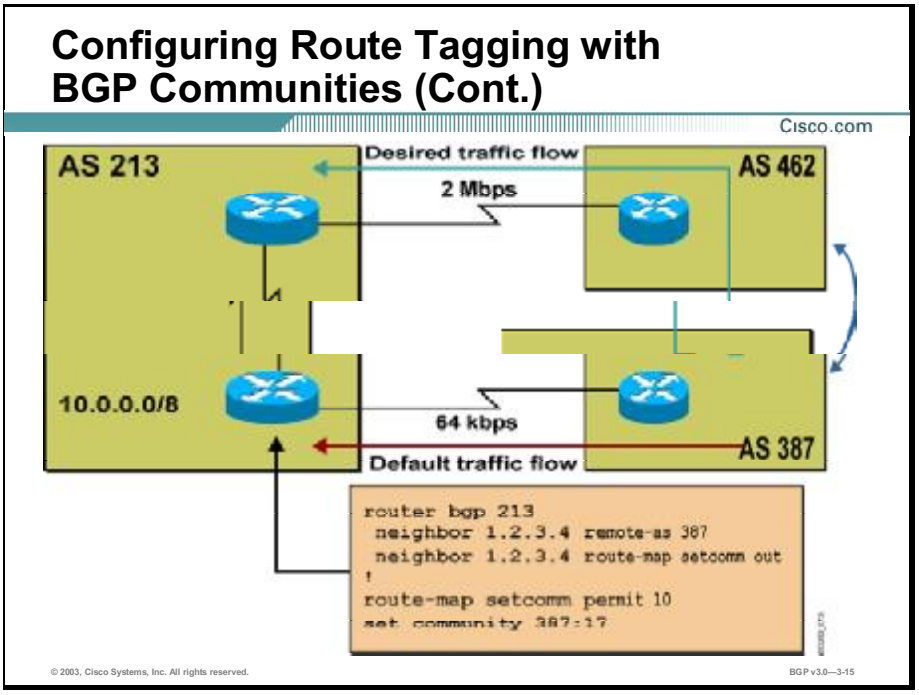
Community: 6553620

ip bgp-community new-format

router# show ip bgp 6.0.0.0

Community: 100:20

Example



In this example, a border router in AS 213 applies a community value 387:17 to all networks sent to neighboring AS 387. In this example, another route-map entry is not needed because the first statement permits all networks (no **match** command means **match all**).

If it is more desirable to set communities on specific routes, you can use a standard access control list (ACL) to match against, with the **match ip address** command in the route-map.

In a later example, networks with community 387:17 will have the local preference changed to a value of 50 within AS 387 to force AS 387 to prefer the other path that carries the default local preference of 100.

Practice

Q1) How many community tags can be attached to a single BGP route?

A) 1

B) 32

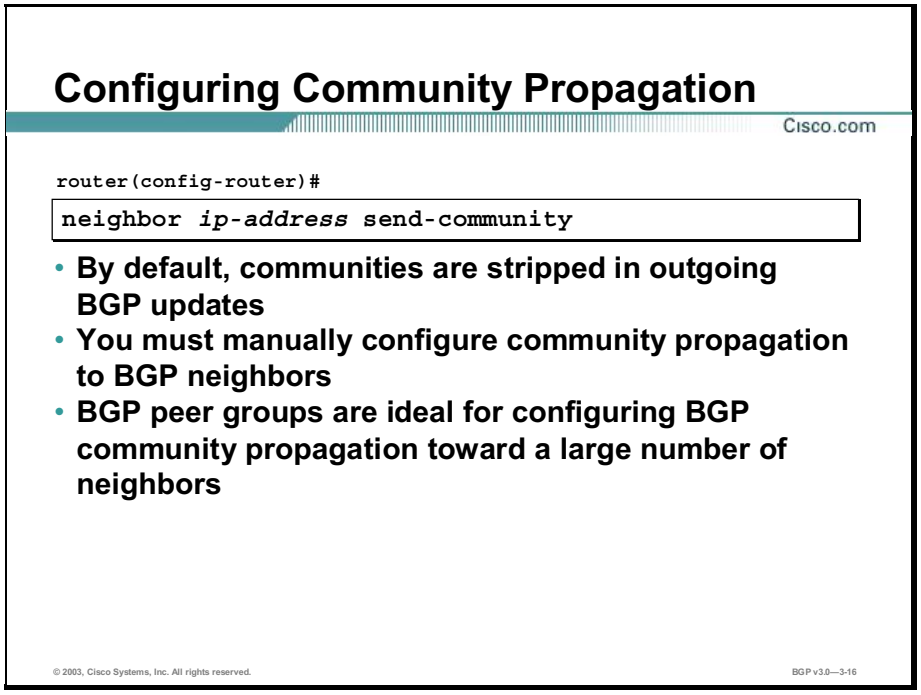
C) 255

D) depends on the number configured with the **ip bgp community** command

 Copyright © 2003, Cisco Systems, Inc.

Configuring Community Propagation

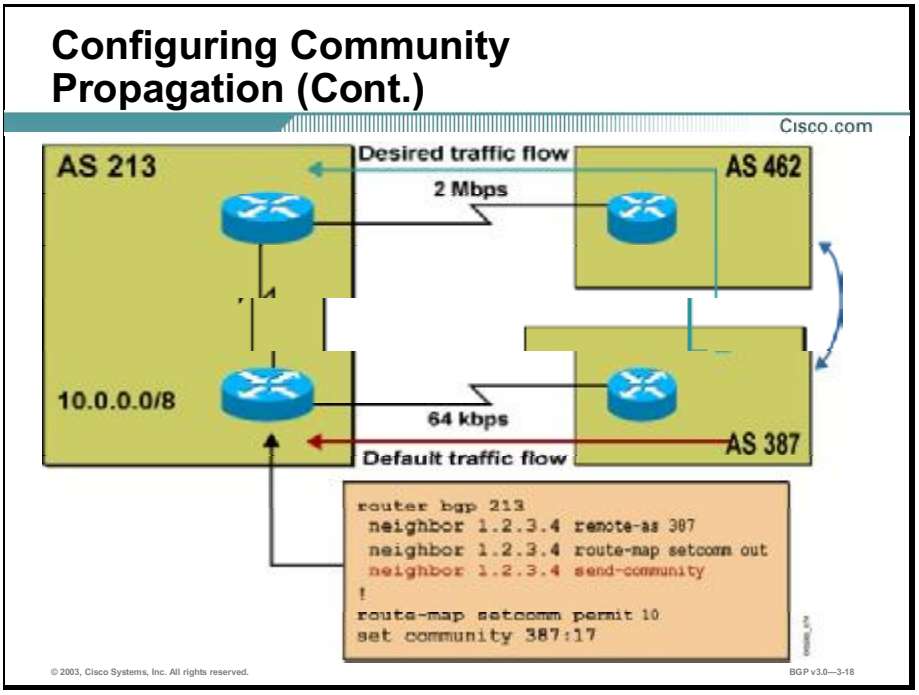
This topic lists the Cisco IOS commands required to enable BGP community propagation to BGP neighbors.


Configuring Community Propagation
Cisco.com
router(config-router)#
neighbor ip-address send-community
• By default, communities are stripped in outgoing BGP updates
• You must manually configure community propagation to BGP neighbors
• BGP peer groups are ideal for configuring BGP community propagation toward a large number of neighbors
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-16

A command commonly forgotten by network administrators when configuring BGP communities is **neighbor** *ip-address* **send-community**. This command is needed to propagate community attributes to BGP neighbors. Even if you use an outgoing route-map to set communities, by default, the router will strip out any community values attached to outgoing BGP updates if you have not configured this command for the specific BGP neighbor.

You can also apply this command to a peer group.

Example



The configuration example discussed earlier in this lesson must include the **send-community** command to enable community propagation from AS 213 to AS 387.

Practice

Q1) What must you configure on a Cisco router to begin the propagation of the community attribute?

A) a BGP peer group

B) **send-community** on the **neighbor** statement

C) a distribute-list under the router BGP process

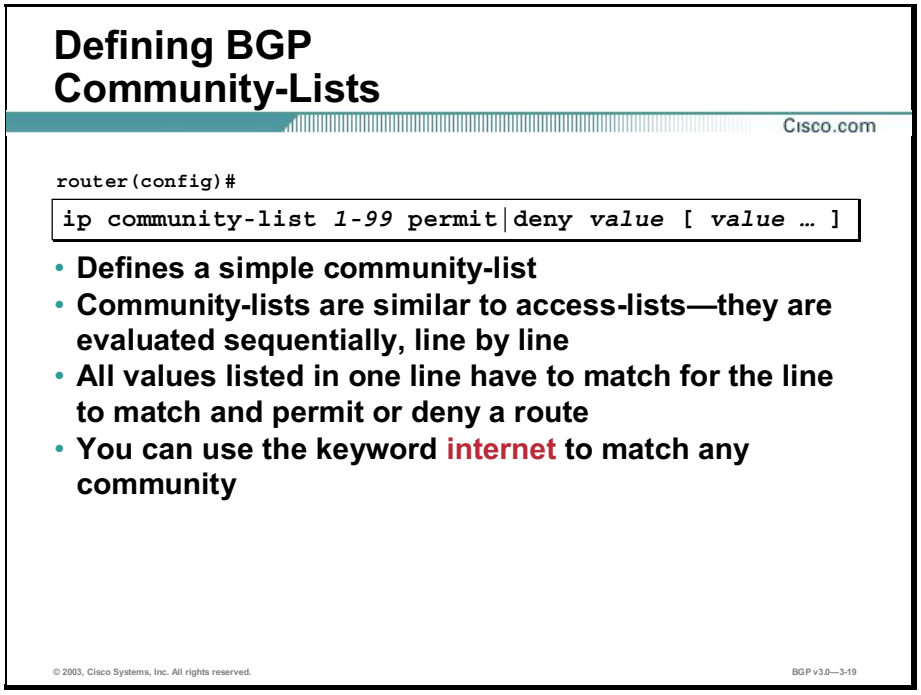
D) an outgoing route-map on the **neighbor** statement

E)

 Copyright © 2003, Cisco Systems, Inc.

Defining BGP Community-Lists

This topic lists the Cisco IOS commands required to match routes based on attached BGP communities using community-lists.

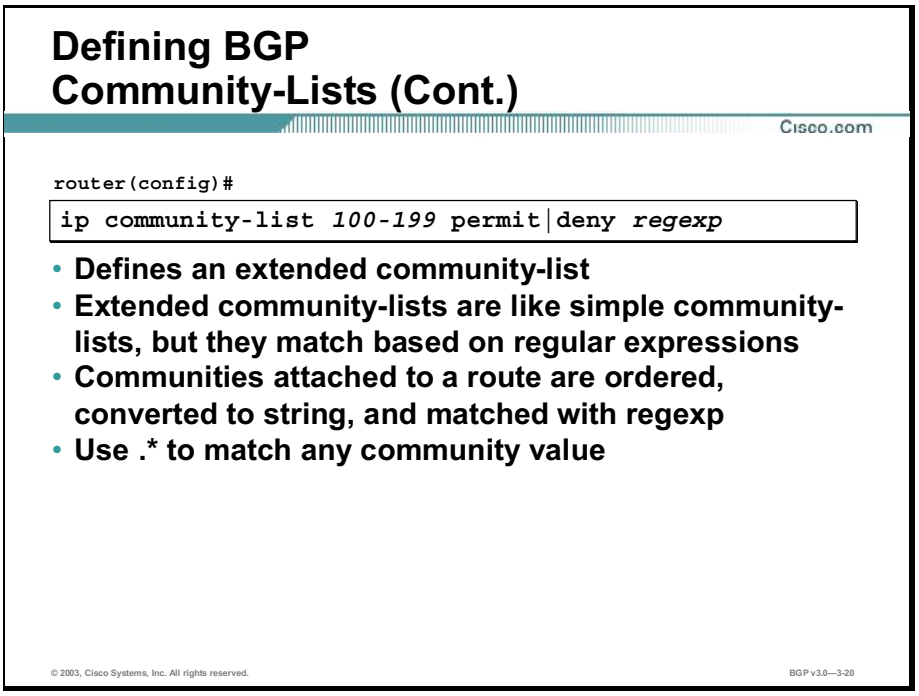

Defining BGP Community-Lists
Cisco.com
router(config)#
ip community-list 1-99 permit|deny value [value …]
• Defines a simple community-list
• Community-lists are similar to access-lists—they are evaluated sequentially, line by line
• All values listed in one line have to match for the line to match and permit or deny a route
• You can use the keyword internet to match any community
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-19

You can use a standard community access-list to find community attributes in routing updates. A standard community-list is defined by its assigned list number, which ranges from 1 to 99. Community-lists are similar to standard IP access–lists in these ways:

- The router evaluates the lines in the community-list sequentially.
- If no line matches communities attached to a BGP route, the route will be implicitly denied.

Standard community-lists are different from standard IP access-lists in these ways:

- The keyword **internet** should be used to permit any community value.
- If more values are listed in a single line, they all have to be in an update to have a match.


Defining BGP
Community-Lists (Cont.)
Cisco.com
router(config)#
ip community-list 100-199 permit|deny regexp
• Defines an extended community-list
• Extended community-lists are like simple community-lists, but they match based on regular expressions
• Communities attached to a route are ordered, converted to string, and matched with regexp
• Use .* to match any community value
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-20

An extended community-list is defined by its assigned list number, which ranges from 100 to 199. Regular expressions are used to match community attributes. When a router processes a list of communities, attached to a network update, they are converted into an ordered string of characters. The example below shows how this is done:

1. The original list of communities in an update:

   “10.0.0.0/8, NH=1.1.1.1, origin=I, AS-path=”20 30 40”, community=10:101, community=10:201, community=10:105, community=10:205”

2. A string of characters containing an ordered list of community values:

   “_10:101_10:105_10:201_10_205_” (“_” represents a space)

3. A regular expression:

   “permit _10:.0[1-5]_” (“_” represents an underscore that matches spaces)

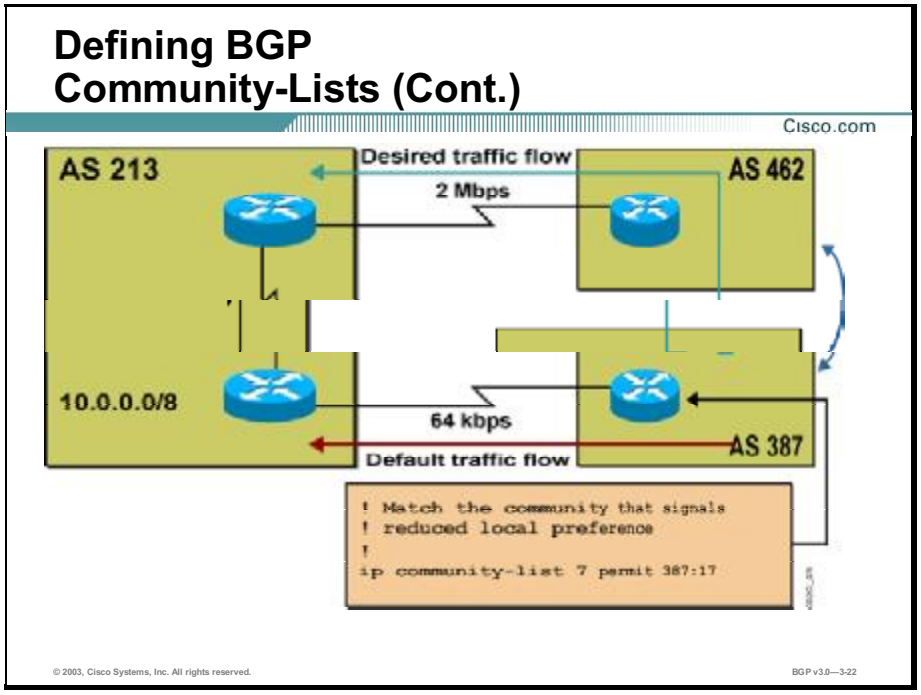
4. The result:

   This regular expression permits the route because it permits all routes carrying communities with the first 16 bits carrying the AS number 10 and the second 16 bits having 0 as the second digit and 1 to 5 as the third digit; the first digit can be anything (as indicated by the “.”).

   Use regular expression “.*” to permit any community.

 Copyright © 2003, Cisco Systems, Inc.

Example



This example shows a portion of the configuration of the router in AS 387. The access-list has been configured to match communities previously set by the router in AS 213.

Practice

Q1) What match criteria are specified in a standard BGP community-list?

A) destination IP addresses

B) regular expressions

C) community attribute values

D) AS numbers

Q2) What regular expression should you use with an extended BGP community-list to match any community value?

A) internet

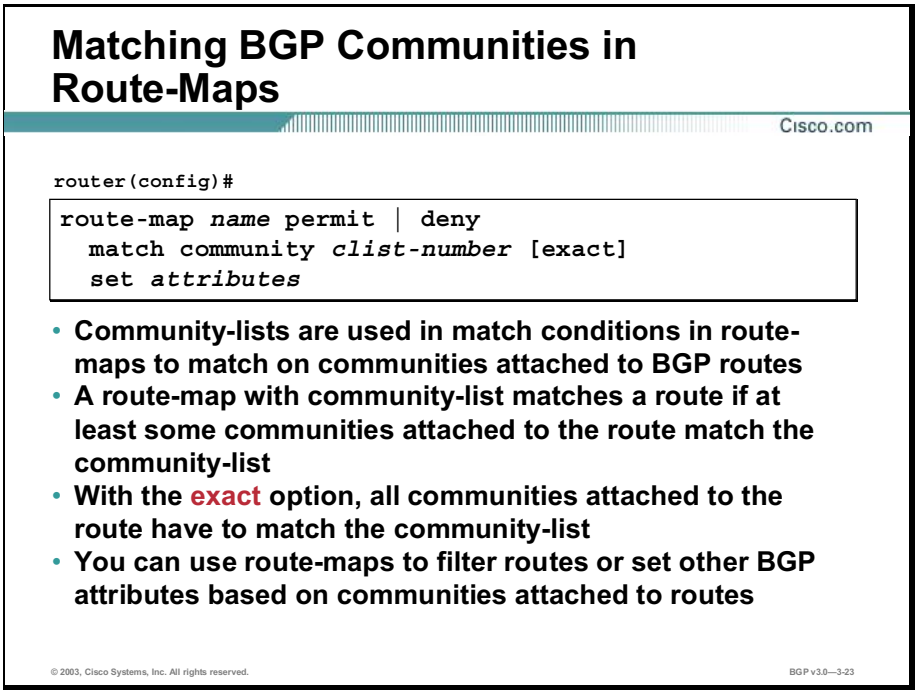
B) any

C) .*

D) permit all

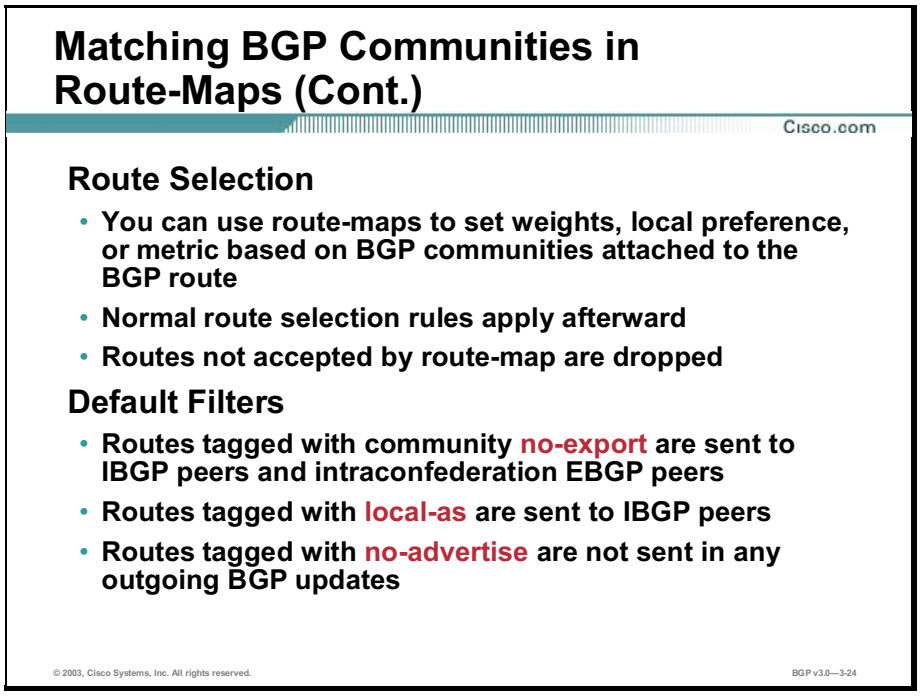
Matching BGP Communities in Route-Maps

This topic lists the commands required to match routes based on attached BGP communities using route-maps.


Matching BGP Communities in Route-Maps
Cisco.com
router(config)#
route-map name permit | deny
match community clist-number [exact]
set attributes
• Community-lists are used in match conditions in route-maps to match on communities attached to BGP routes
• A route-map with community-list matches a route if at least some communities attached to the route match the community-list
• With the exact option, all communities attached to the route have to match the community-list
• You can use route-maps to filter routes or set other BGP attributes based on communities attached to routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-23

Network administrators use route-maps to match networks that carry a subset of communities that are permitted by the community-list. Other parameters or attributes can then be set based on community values. If you use the keyword **exact**, all communities attached to a BGP route have to be matched by the community-list.

 Copyright © 2003, Cisco Systems, Inc.

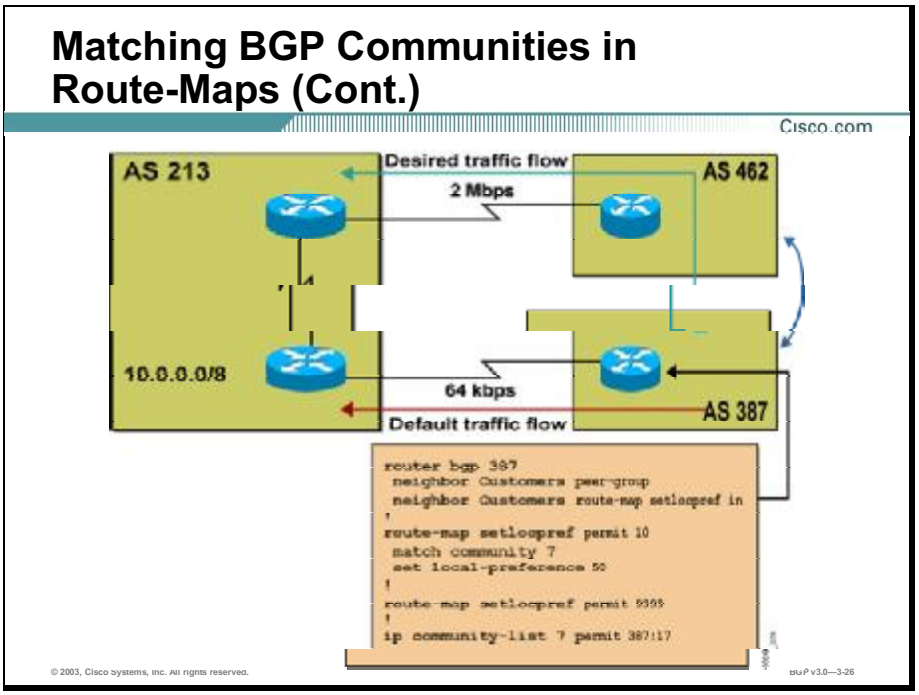

Matching BGP Communities in Route-Maps (Cont.)
Cisco.com
Route Selection
• You can use route-maps to set weights, local preference, or metric based on BGP communities attached to the BGP route
• Normal route selection rules apply afterward
• Routes not accepted by route-map are dropped
Default Filters
• Routes tagged with community no-export are sent to IBGP peers and intraconfederation EBGP peers
• Routes tagged with local-as are sent to IBGP peers
• Routes tagged with no-advertise are not sent in any outgoing BGP updates
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-24

You can use a route-map to filter or modify BGP routing updates. Any BGP-related **set** commands can be used to set BGP parameters and attributes (that is, weight, local preference, multi-exit discriminator [MED]).

As mentioned before, there are some predefined community values that cause routers to automatically filter routing updates:

- **no-export:** If a router receives an update carrying this community, it will not propagate it to any external neighbors except to intraconfederation external neighbors.
- **no-advertise:** If a router receives an update carrying this community, it will not forward it to any neighbor.
- **local-as:** This community has a similar meaning to **no-export**, but it keeps a route within the local subautonomous system. It is not sent to intraconfederation external neighbors or to any other external neighbors.
- **internet:** Advertise this route to the Internet community. All routers belong to it.

Example



This example shows a configuration that translates community 387:17 into local preference 50. All updates received from neighboring AS 213 are processed by the route-map, which uses a community-list to find community 387:17. If the community-list matches one of the community attributes, the **set** command is executed and the route is permitted. If the route does not contain the right community, it is simply permitted by route-map statement 9999 without changing anything in the update.

The result is that AS 387 prefers other paths to AS 213 because they have a default local preference of 100.

Practice

Q1) What is the result of tagging a route with the no-export community?

A) The route will not be advertised within the local AS.

B) The upstream AS will not be allowed to export the route.

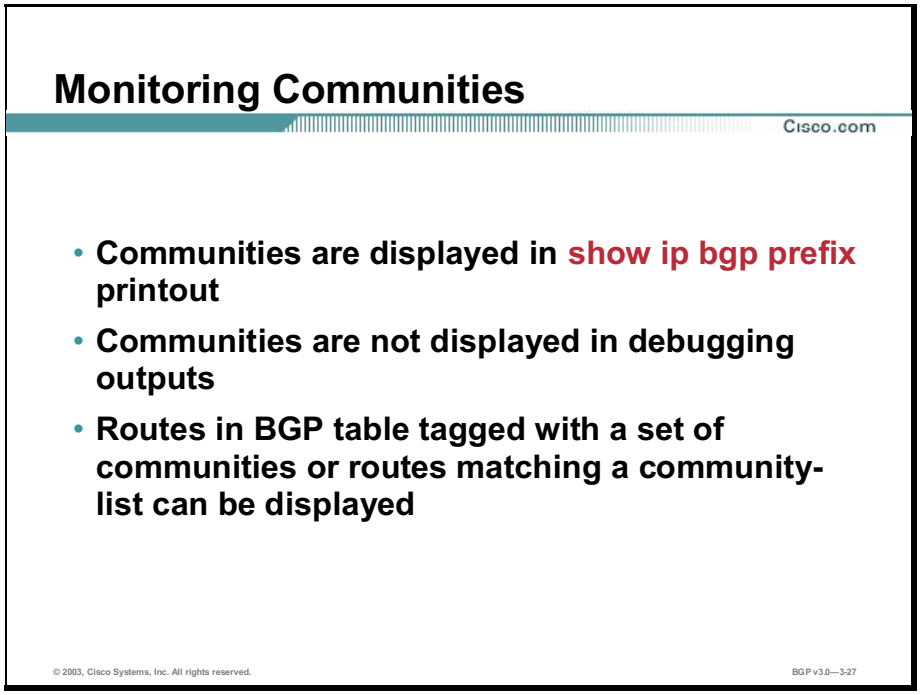
C) The route cannot be exported to another routing protocol.

D) The router will not propagate the route to any external neighbors except to intraconfederation external neighbors.

 Copyright © 2003, Cisco Systems, Inc.

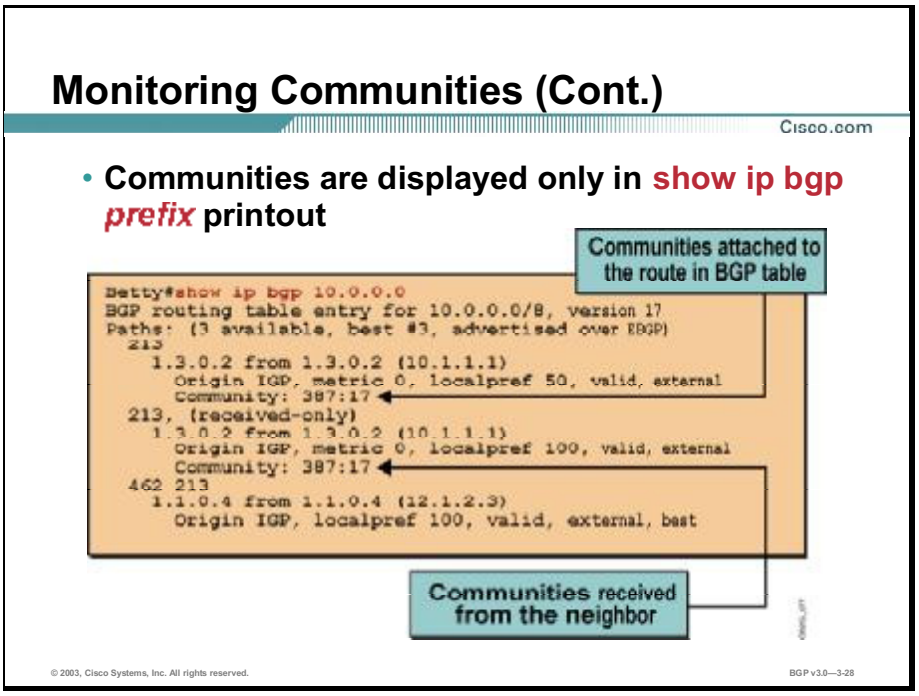
Monitoring Communities

This topic lists the commands required to monitor BGP communities.


Monitoring Communities
Cisco.com
• Communities are displayed in show ip bgp prefix printout
• Communities are not displayed in debugging outputs
• Routes in BGP table tagged with a set of communities or routes matching a community-list can be displayed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-27

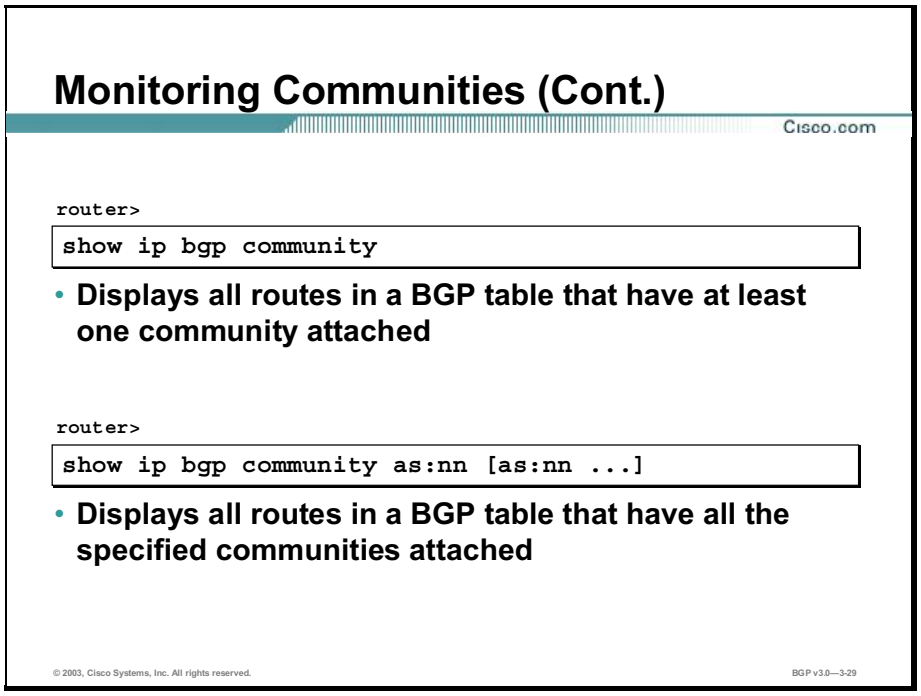
Because a community is an attribute that can appear more than once in a single update, the **show ip bgp** command does not show it. You can view communities only if you use the **show ip bgp** *prefix* command.

If you use the **show ip bgp community-list** command, all networks are listed that are permitted by the community-list.


Monitoring Communities (Cont.)
Cisco.com
• Communities are displayed only in show ip bgp prefix printout
Communities attached to the route in BGP table
Betty#show ip bgp 10.0.0.0
BGP routing table entry for 10.0.0.0/8, version 17
Paths: (3 available, best #3, advertised over EBGP)
213
1.3.0.2 from 1.3.0.2 (10.1.1.1)
Origin IGP, metric 0, localpref 50, valid, external
Community: 387:17
213, (received-only)
1.3.0.2 from 1.3.0.2 (10.1.1.1)
Origin IGP, metric 0, localpref 100, valid, external
Community: 387:17
462 213
1.1.0.4 from 1.1.0.4 (12.1.2.3)
Origin IGP, localpref 100, valid, external, best
Communities received from the neighbor
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-28

This example shows the output of the **show ip bgp** *prefix* command where inbound soft reconfiguration was enabled on one of the neighbors. The original update contained one single community attribute (387:17), which can be seen from the second path marked with received-only. This update was then processed by an inbound route map, which matched the community 387:17 and changed the local preference of the received route to 50.

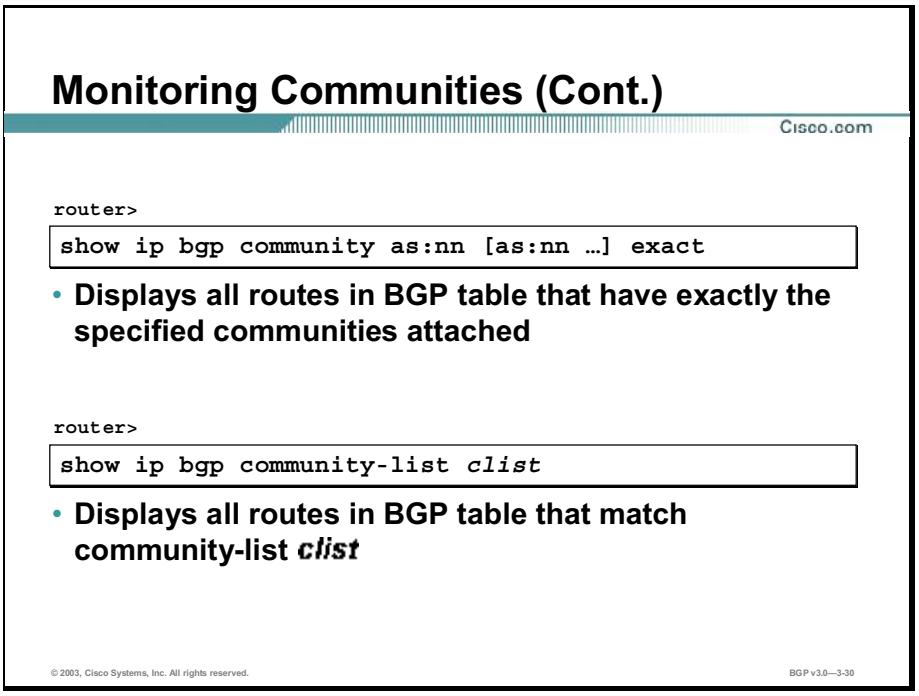
 Copyright © 2003, Cisco Systems, Inc.


Monitoring Communities (Cont.)
Cisco.com
router>
show ip bgp community
• Displays all routes in a BGP table that have at least one community attached
router>
show ip bgp community as:nn [as:nn ...]
• Displays all routes in a BGP table that have all the specified communities attached
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-29

Another use of the **show** command is to filter the output of **show ip bgp**.

If the keyword **community** is included, all networks that have at least one community attribute are displayed.

If the keyword **community** is followed by one or more community values, only those networks that carry all those communities are displayed.


Monitoring Communities (Cont.)
Cisco.com
router>
show ip bgp community as:nn [as:nn …] exact
• Displays all routes in BGP table that have exactly the specified communities attached
router>
show ip bgp community-list clist
• Displays all routes in BGP table that match community-list clist
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-30

If the keyword **community** is followed by one or more community values, only those networks that carry all those communities are displayed. If the keyword **exact** is added at the end, only those that match exactly are displayed.

You can also use a community-list to filter the output of the **show ip bgp** command.

Practice

Q1) What command should you use to display the community attribute?

A) **show ip bgp summary**

B) **show community**

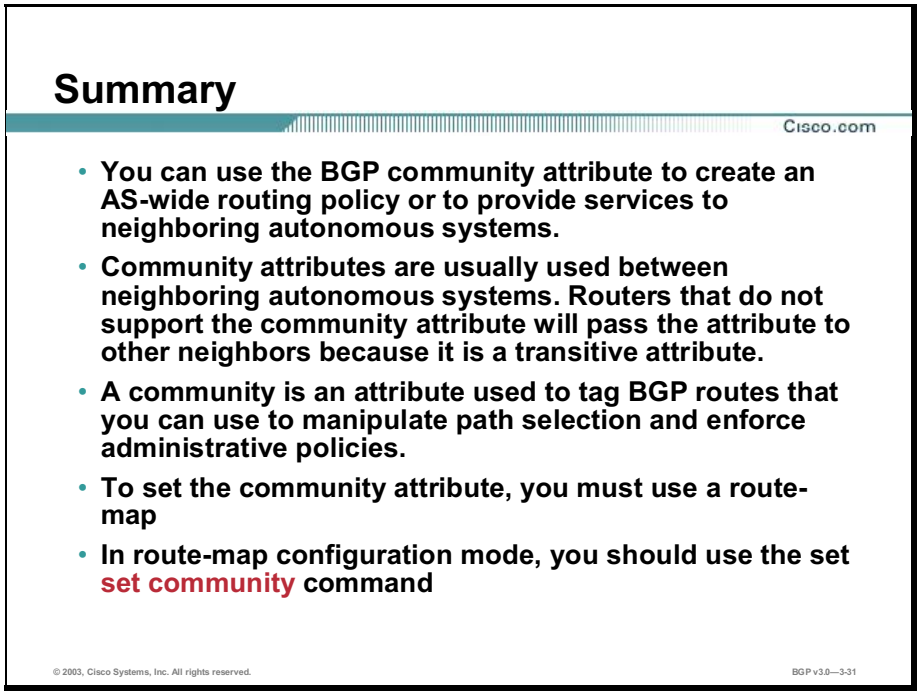
C) **show ip bgp** *prefix*

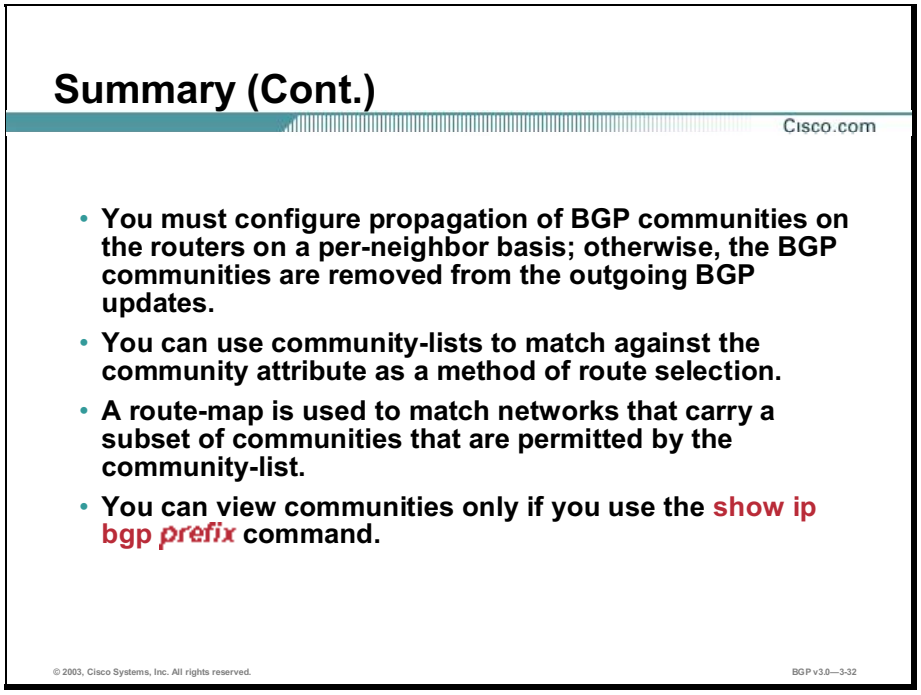
D) **show ip route**

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• You can use the BGP community attribute to create an AS-wide routing policy or to provide services to neighboring autonomous systems.
• Community attributes are usually used between neighboring autonomous systems. Routers that do not support the community attribute will pass the attribute to other neighbors because it is a transitive attribute.
• A community is an attribute used to tag BGP routes that you can use to manipulate path selection and enforce administrative policies.
• To set the community attribute, you must use a route-map
• In route-map configuration mode, you should use the set set community command
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-31


Summary (Cont.)
Cisco.com
• You must configure propagation of BGP communities on the routers on a per-neighbor basis; otherwise, the BGP communities are removed from the outgoing BGP updates.
• You can use community-lists to match against the community attribute as a method of route selection.
• A route-map is used to match networks that carry a subset of communities that are permitted by the community-list.
• You can view communities only if you use the show ip bgp prefix command.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—3-32

Next Steps

After completing this lesson, go to:

- Customer-to-Provider Connectivity with BGP module

References

For additional information, refer to these resources:

- For more information on BGP communities, refer to “BGP Case Studies” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#communityattribute

- For further information on BGP communities, refer to “Configuring BGP” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/np1_c/1cprt1/1cbgp.htm#xtocid34

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: BGP Communities

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using the BGP community attribute in a situation where you must support multiple connections to an Internet service provider (ISP).

After completing this exercise, you will be able to:

- Configure route tagging using BGP communities
- Configure BGP community propagation
- Monitor BGP communities

Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in | out**}	Use this command to apply a route-map to incoming or outgoing routing updates.
route-map *name* {**permit | deny**} **seq**	Define or modify an existing entry in a route-map.
match community *c-acl*	Use this command in a route-map to match communities by using a community-list.
set community *community* [**additive**]	Use this command in a route-map to set community attributes. Use the keyword **additive** to append communities instead of replacing them.
set local-preference *num*	Use this command to set local preference attribute.
ip bgp-community new-format	Configure BGP communities using AA:NN format.
ip community-list *num* {**permit|deny**} **community**	Use this command to create a community access-list.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of **show ip bgp** command.
show ip bgp community [**community** [**community** …]] [**exact-match**]	Use this command to view BGP routes that have at least one community attribute or those specified in the command.
show ip bgp community-list *c-list* [**exact-match**]	Use this command to view BGP routes that are permitted by the specified community-list.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- Your second service provider “Cheap” has upgraded its link to your network to match that of the provider “Good.” You decide to change your route selection policy to use the newly updated link from “Cheap.” As a result of the new service available from “Cheap,” you decide to remove the backup link to provider “Good” as you can now achieve both link and provider redundancy.

- You wish for all outgoing traffic to still use the link to “Good,” but all return traffic should use the newly updated link from “Cheap.” Knowing that you have a connection to the provider “Good,” you still wish to use the primary connection to “Good” for return traffic should the link to provider “Cheap” fail.

 Copyright © 2003, Cisco Systems, Inc.

- You cannot influence return path selection across multiple providers and solicit assistance from provider “Good.” The provider “Good” agrees to set local preference for any routes that you mark as follows:
  - Local preference of 50 for routes marked with a community attribute of x:20
- The provider has preconfigured router “Good” to set local preference on IP routes set with the correct community attribute to assist you in influencing return path selection.
- Figure 1 shows the physical connectivity, BGP sessions, and traffic flow in the network.

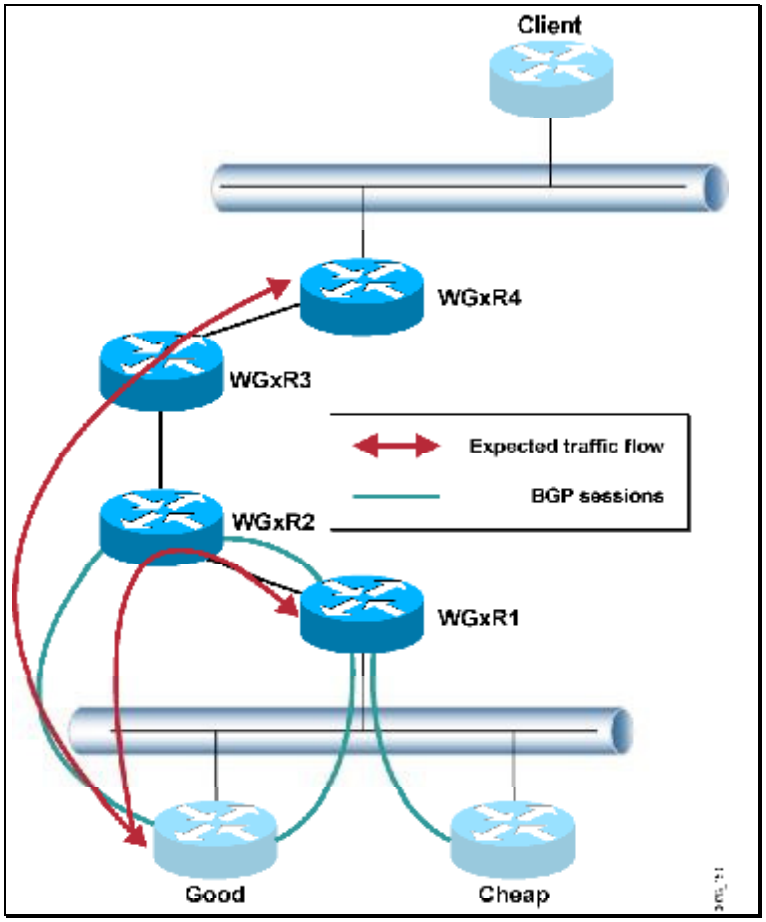


Figure 1: BGP communities physical connectivity

Exercise Procedure

Complete these steps:

Step 1 Before starting this lab, use a trace from router "Good" to 197.x.7.1. The trace should match that below (where x is your WG number).

```
Good#traceroute 197.x.7.1

Type escape sequence to abort.
Tracing the route to 197.x.7.1

1 192.168.3x.1 [AS 1] 16 msec 16 msec 16 msec
2 192.168.x.6 [AS 1] 28 msec 32 msec 28 msec
3 192.168.x.10 [AS 1] 44 msec *  40 msec
```

Step 2 Remove the BGP neighbor statement on WGxR1 for router "Good."

Configure BGP community propagation:

Step 3 Configure BGP community propagation on router WGxR2.

Step 4 Enable your WGxR2 router to configure communities using the AA:NN format.

Configure outbound filters on WGxR2:

Step 5 Create a route-map to set the community attribute on all BGP updates sent from WGxR2 to the provider "Good." Set the community attribute to x:20, where x is your workgroup number.

Step 6 Apply the route-map to the BGP neighbor "Good" in the outbound direction.

Exercise Verification

You have completed this exercise when you attain these results:

- Log on to the "Good" router and verify that the BGP community attribute has been correctly set.

```
Good>show ip bgp community 1:20

Good#sh ip bgp community 1:20
BGP table version is 61, local router ID is 199.199.199.199
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  192.168.1.0      192.168.31.1             0     50      0 1 i
*  192.168.31.0     192.168.31.1                   50      0 1 i
*  197.1.0.0/21     192.168.31.1                   50      0 1 i
*  197.1.0.0/16     192.168.31.1                   50      0 1 i
*  197.1.8.0/22     192.168.31.1                   50      0 1 i
```

 Copyright © 2003, Cisco Systems, Inc.

- Log in to router “Good” and perform the same traceroute that was done before starting the configuration in this lab.

```
Good#traceroute 197.1.7.1
Type escape sequence to abort.
Tracing the route to 197.1.7.1

  1 wg1 (192.168.20.1) 0 msec 0 msec 0 msec
  2 192.168.1.2 [AS 1] 16 msec 16 msec 16 msec
  3 192.168.1.6 [AS 1] 28 msec 28 msec 36 msec
  4 192.168.1.10 [AS 1] 44 msec *  44 msec
```

Note Normally, you would expect to see the first line of the traceroute showing “Cheap” as the first hop. However, because the routers “Good,” “Cheap,” and WGxR1 share a common Ethernet segment, the BGP next-hop rule states: If the current BGP next hop is in the same IP subnet as the receiving router, the next hop is not changed; otherwise, it is changed to the IP address of the sending router. The next-hop rule applies even when there is no BGP connection between “Good” and WGxR1.

- On router “Good,” execute the command: **show ip bgp 197.1.7.1**.

```
Good#sh ip bgp 197.1.7.1
BGP routing table entry for 197.1.0.0/21, version 22
Paths: (2 available, best #2, table Default-IP-Routing-Table)
  Advertised to peer-groups:
     students2
  1, (aggregated by 1 197.1.8.1)
    192.168.31.1 from 192.168.31.1 (197.1.3.1)
      Origin IGP, localpref 50, valid, external, atomic-aggregate
      Community: 1:20

  22 1, (aggregated by 1 197.1.8.1)
    192.168.20.1 from 192.168.20.22 (192.20.11.1)
      Origin IGP, localpref 100, valid, external, atomic-aggregate,
best
```

Note that the best path, because of local preference, is the bottom route, and the next hop is 192.168.20.22 – “Cheap” – but because of the shared media, the next hop has not been changed. This result is true even though the AS-path length is longer, because local preference is a stronger route selection tool.

Answer these questions:

Q1) What do you have to do to enable community propagation?

Q2) What mechanisms can you use to match or set communities?

 Copyright © 2003, Cisco Systems, Inc.

4

Customer-to-Provider Connectivity with BGP

Overview

Today, many companies use the Internet for a variety of reasons, including increasing employee productivity, increasing sales, increasing customer satisfaction, and reducing cycle time. A key component in connecting companies to the Internet is the service provider. Depending upon business goals, application requirements, and administrative policies, a company will use different methods to connect to a service provider. Company business requirements and policies may even dictate that the company connects to multiple service providers. This module discusses the different requirements for connectivity between customers and service providers. Included in this module is a discussion of physical connection methods, redundancy, load balancing, and technical requirements such as addressing and autonomous system (AS) numbering. This module discusses the configuration requirements to connect a customer to a single service provider using static routes and using the Border Gateway Protocol (BGP). Also discussed in this module are the configuration requirements to connect a customer to multiple service providers using BGP.

Upon completing this module, you will be able to:

- Describe the requirements to connect customer networks to the Internet in a service provider environment
- Implement customer connectivity using static routing, given a service provider network
- Implement customer connectivity using BGP, given a customer scenario where you must support multiple connections to a single ISP
- Implement customer connectivity using BGP, given a customer scenario where you must support connections to multiple ISPs

Outline

The module contains these lessons:

- Customer-to-Provider Connectivity Requirements
- Implementing Customer Connectivity Using Static Routing
- Connecting a Multihomed Customer to a Single Service Provider
- Connecting a Multihomed Customer to Multiple Service Providers

 Copyright © 2003, Cisco Systems, Inc.

Customer-to-Provider Connectivity Requirements

Overview

Customers connect to the Internet using service providers to enable different applications such as intranet connectivity with Virtual Private Networks (VPNs), extranet connectivity with suppliers, and other Internet applications. Different customers have different connectivity requirements depending upon their business model, redundancy requirements, and even network budget.

This lesson discusses different solutions for connecting customer networks to service providers. Included in this lesson is a discussion of customer network redundancy requirements, routing requirements, IP addressing requirements, and autonomous system (AS) numbering requirements.

Importance

When planning network connectivity to a service provider, network designers must give careful consideration to the different aspects of the connectivity, including physical connection types, redundancy provided by the connection method chosen, IP addressing requirements, and AS numbering considerations, if the network design is going to meet both the business and technical requirements of the applications planned for the network.

Objectives

Upon completing this lesson, you will be able to:

- Identify different physical connections used by customers to connect to a service provider
- Describe the levels of redundancy provided by each physical connection type used by customers to connect to a service provider

- Identify different routing schemes used by customers to connect to a service provider
- Describe routing schemes that are appropriate for each physical connection type used by customers to connect to a service provider
- Describe the addressing schemes used by customers to connect to a service provider
- Describe AS numbering schemes used by customers to connect to a service provider

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

Outline

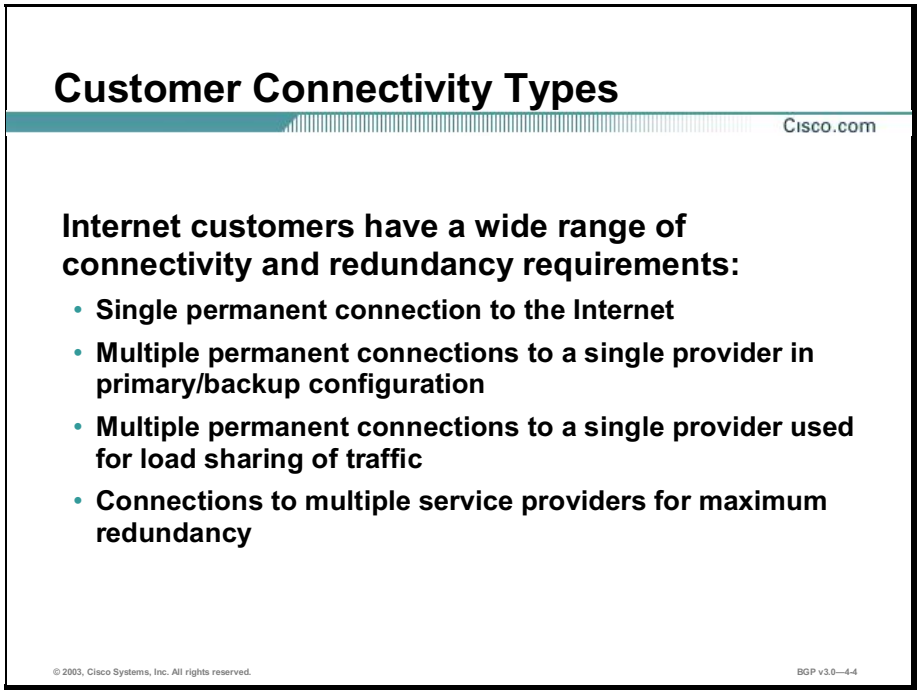
This lesson includes these topics:

- Overview
- Customer Connectivity Types
- Redundancy in Customer Connections
- Customer-to-Provider Routing Schemes
- Customer Routing
- Addressing Requirements
- AS Number Allocation
- Summary
- Assessment (Quiz): Customer-to-Provider Connectivity Requirements

 Copyright © 2003, Cisco Systems, Inc.

Customer Connectivity Types

This topic identifies different physical connections used by customers to connect to a service provider.


Customer Connectivity Types
Cisco.com
Internet customers have a wide range of connectivity and redundancy requirements:
• Single permanent connection to the Internet
• Multiple permanent connections to a single provider in primary/backup configuration
• Multiple permanent connections to a single provider used for load sharing of traffic
• Connections to multiple service providers for maximum redundancy
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-4

Service provider customers have different requirements for their Internet connectivity. These different requirements result in different solutions:

- A single permanent connection to one Internet service provider (ISP). This solution meets the requirements for the vast majority of customers.
- Multiple permanent connections where one of the lines is primary and the other line is used for backup only. This setup also provides redundancy on the links. Compared to a dial-up backup, a permanent backup link is preferred for various reasons such as the severe bandwidth limitations on dial-up lines and the time required to establish a dial-up connection.
- Multiple permanent connections to one ISP used for load sharing of traffic. This solution gives redundancy on the links but also provides additional bandwidth.
- Permanent connections to more than one ISP. This solution provides the highest level of redundancy, because it not only can cope with link-level failures but also failures within the network of a service provider.

Practice

Q1) Which method of Internet connectivity gives the highest level of redundancy?

A) a single permanent connection to one ISP

B) multiple permanent connections to one ISP where one of the lines is primary and the other line is used for backup only

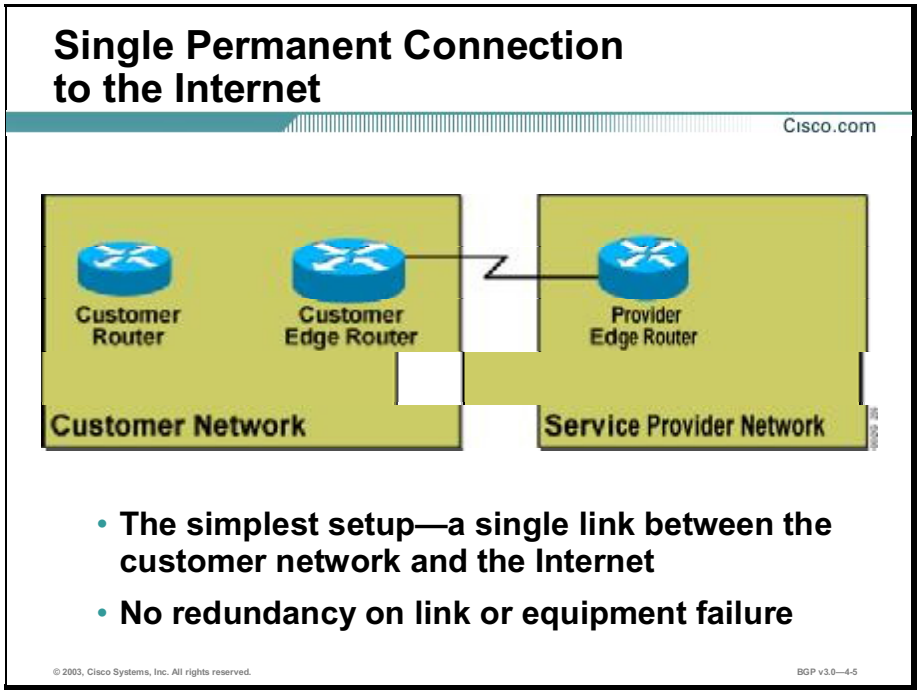
C) multiple permanent connections to one ISP used for load sharing of traffic

D) multiple permanent connections to more than one ISP

 Copyright © 2003, Cisco Systems, Inc.

Redundancy in Customer Connections

This topic describes the different levels of redundancy provided by each physical connection type used by customers to connect to a service provider.

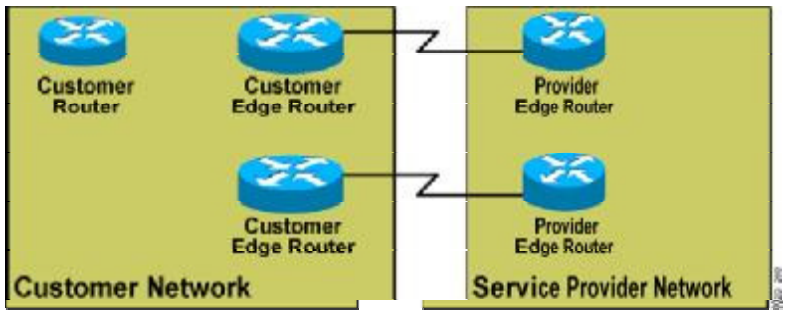

Single Permanent Connection to the Internet
Cisco.com
Customer Router
Customer Edge Router
Provider Edge Router
Customer Network
Service Provider Network
• The simplest setup—a single link between the customer network and the Internet
• No redundancy on link or equipment failure
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-5

A single permanent connection to one ISP is the most common setup. This setup is also the simplest to implement.

The customer network has an edge router. This router is connected to one of the edge routers of the ISP. The connection is permanent and could be a leased line, a Frame Relay or ATM permanent virtual connection (PVC), a LAN segment, or something equivalent.

There is no redundancy in this solution. Any failure on the permanent link or any of the two edge routers causes a complete outage of the service. Serious failures within the ISP network, which affect all customers of this ISP, also affect the customer in this example.

Multiple Permanent Connections Providing Redundancy



- **Customers wanting increased redundancy install several physical links to the Internet**
- **Redundant link is used in primary/backup setup or for load sharing**
- **Redundancy on link or equipment failure**
- **No redundancy on service provider failure**

© 2003, Cisco Systems, Inc. All rights reserved.

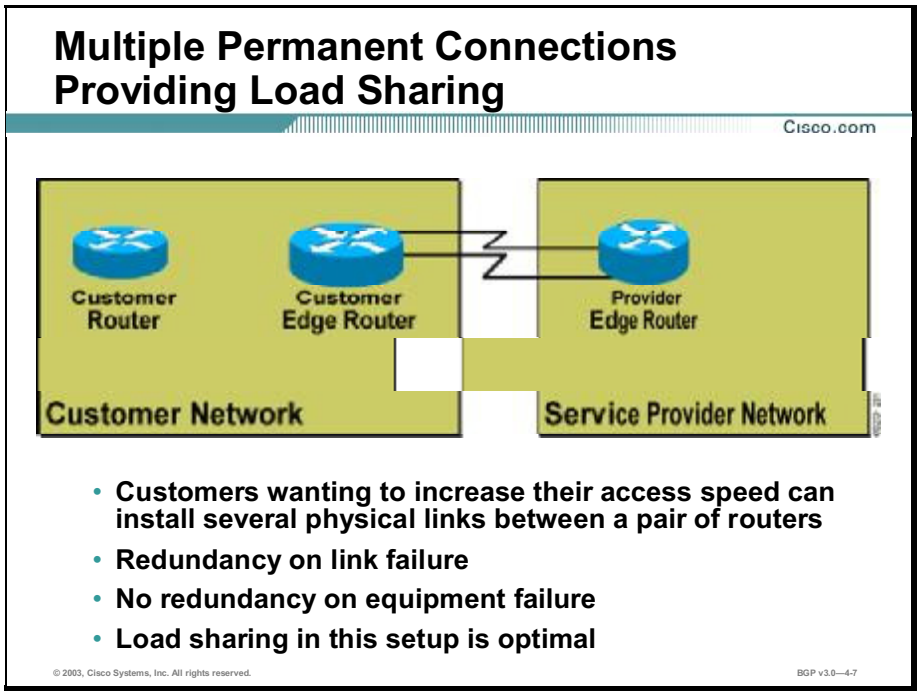
In this setup, one customer edge router connects to one ISP edge router. A different customer edge router is used to connect to another ISP edge router. If one of these routers fails, only one of the connections breaks down. The other connection is still available.

In some cases, the two links may be implemented between the customer and the provider for load sharing and in other cases strictly for backup purposes. For example, backup PVCs in Frame Relay or ATM networks can sometimes be very cost-efficient, provided that these PVCs carry only a very small volume of traffic and that the primary path is available.

In a case where load sharing between both links is a desired network characteristic, the distribution of the load over the links is more complicated compared to a case where both links terminate in the same router.

Again, because the customer is connected to a single ISP, serious ISP network failures, which affect all customers of this ISP, will also affect the customer in this scenario, regardless of the backup link.

 Copyright © 2003, Cisco Systems, Inc.

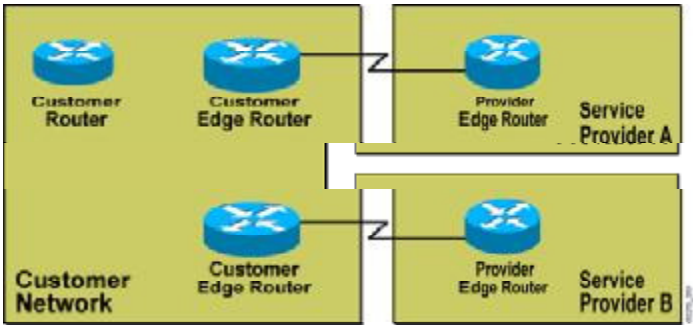


In the example shown, a single router in the customer network is connected to a single router in the ISP network. The redundancy is limited to the link level because router failures are not covered. Using two parallel links between two routers allows for an optimal distribution of load over the links.

Depending on the switching path used in the customer and the ISP routers, load sharing can be performed based on the destination address only (fast switching), based on source-destination address pairs (default behavior for Cisco Express Forwarding [CEF]), or on a packet-by-packet basis (process switching or CEF).

As in the previous examples, serious ISP network failures that affect all customers of this ISP will also affect this customer, regardless of the link backup.

Connections to Multiple Service Providers



- **Customers with maximum redundancy requirements install physical links to multiple Internet service providers**
- **Redundancy on link, equipment or service provider failure**
- **Primary/backup setup is complex without service provider assistance**
- **Good load sharing is impossible to achieve**

© 2003, Cisco Systems, Inc. All rights reserved.

In the example shown, two different edge routers in the customer network have one permanent connection each to different ISPs. Link failures and router failures are covered by the redundancy in exactly the same way as in the example where the two customer routers are connected to two different routers in one ISP network. However, because the two connections in this example go to two different ISPs, the redundancy also covers problems within one ISP network.

The two links may in some cases be implemented by the customer for load sharing and in other cases be used strictly for backup purposes. Controlling load distribution over the links is more complicated in this example. Avoiding any load on the backup link may require assistance from the ISP to which the backup link is connected.

Load sharing between the links in this setup can never be optimal. Equal distribution of the return traffic load from the Internet over the two separate links cannot be done. Distribution of the load of outgoing traffic is done based on destination addresses. Slowly adjusting the appropriate router configuration parameters and observing the link traffic load changes that result, can be used to reach an acceptable distribution of router traffic between the two links.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Multiple permanent connections to a single ISP do not provide what level of redundancy?

A) redundancy on link failure

B) redundancy on ISP failure

C) redundancy on equipment failure

D) redundancy on routing protocol failure

Q2) What is a drawback when a customer is installing multiple permanent connections to multiple ISPs?

A) There is no redundancy on ISP failure.

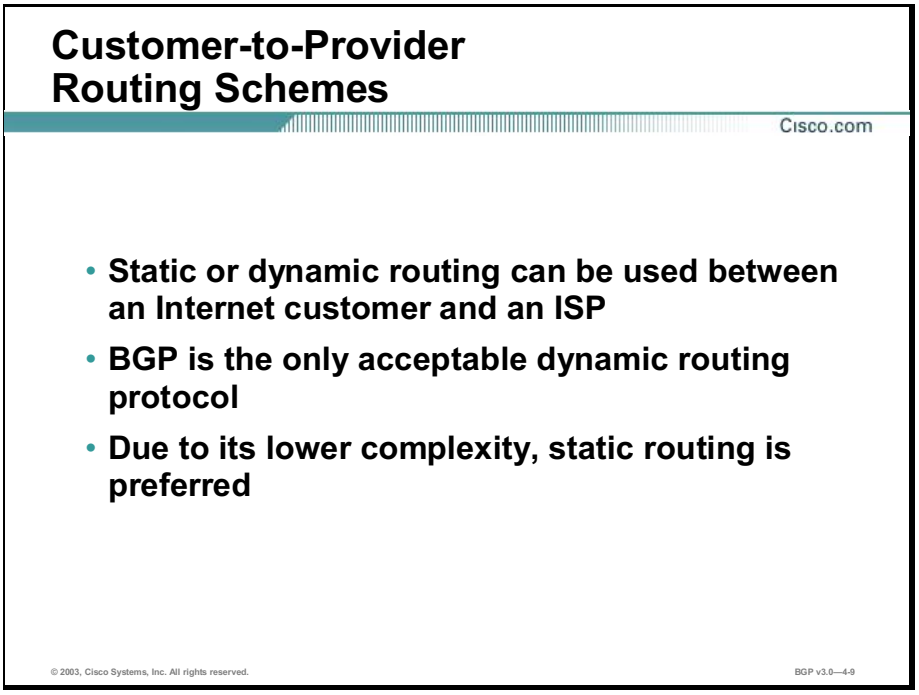
B) Good load sharing is impossible to achieve.

C) The customer can use only Frame Relay PVCs.

D) Equipment failure may cause a complete network outage.

Customer-to-Provider Routing Schemes

This topic identifies different routing schemes that customers use to connect to a service provider.


Customer-to-Provider
Routing Schemes
Cisco.com
• Static or dynamic routing can be used between an Internet customer and an ISP
• BGP is the only acceptable dynamic routing protocol
• Due to its lower complexity, static routing is preferred
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-9

Different solutions for connecting a customer network to the network of an ISP require different methods of routing information exchange.

- Static routing is preferred due to its lower complexity. In a normal case, the customer network must have a default route to the ISP network and the ISP network must have a route to those IP prefixes that the customer has in its network. As always, static routing provides very low, if any, redundancy.

- Dynamic routing provides redundancy. The customer and the ISP networks must be configured to exchange a common routing protocol. BGP is the only choice due to the large volumes of routing information, the inherent security mechanisms of BGP, and the ability of BGP to handle routing policies.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Which form of routing provides the best redundancy?

A) static routing

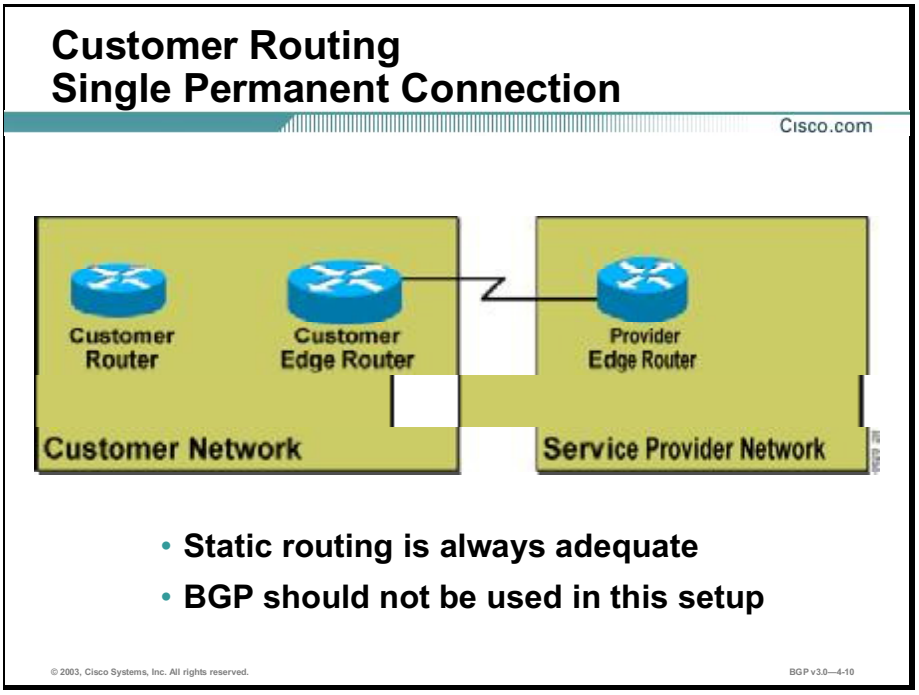
B) content routing

C) dynamic routing

D) embryonic routing

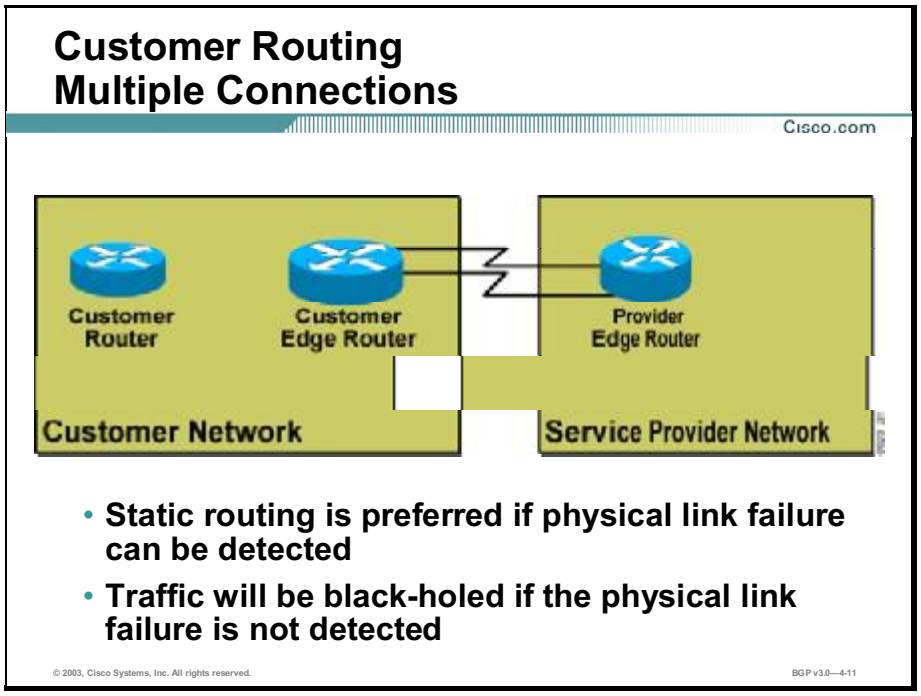
Customer Routing

This topic describes routing schemes that are appropriate for each physical connection type used by customers to connect to a service provider.



In a case where the customer has a single permanent connection to the Internet, static routing is always adequate. The physical topology does not provide any redundancy, and it is therefore unnecessary to add the complexity of dynamic routing. Keep the network simple by avoiding the use of BGP in this case.

 Copyright © 2003, Cisco Systems, Inc.

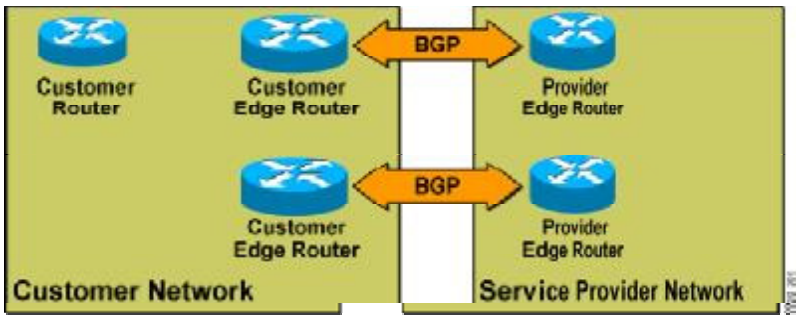

Customer Routing
Multiple Connections
Cisco.com
Customer Router
Customer Edge Router
Provider Edge Router
Customer Network
Service Provider Network
• Static routing is preferred if physical link failure can be detected
• Traffic will be black-holed if the physical link failure is not detected
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-11

Multiple permanent connections between a single router on the customer network and a single router on the service provider network should be configured with static routing, provided that link failure can be detected by link-level procedures.

With this type of connection, two static routes are configured on each network, pointing to both links between the customer and the ISP. If either of the links fails, the link-level procedures should detect this failure and place the interface in a down state. In this case, the static route is invalid and not used for forwarding packets. The router will subsequently forward all packets over the remaining link.

If the link-level procedures cannot detect a link failure, the static route pointing out over the failed link is still valid. The router continues using this static route to send some of the traffic out on the failed interface. This effectively creates a black hole for some of the traffic.

Customer Routing Multiple Connections (Cont.)



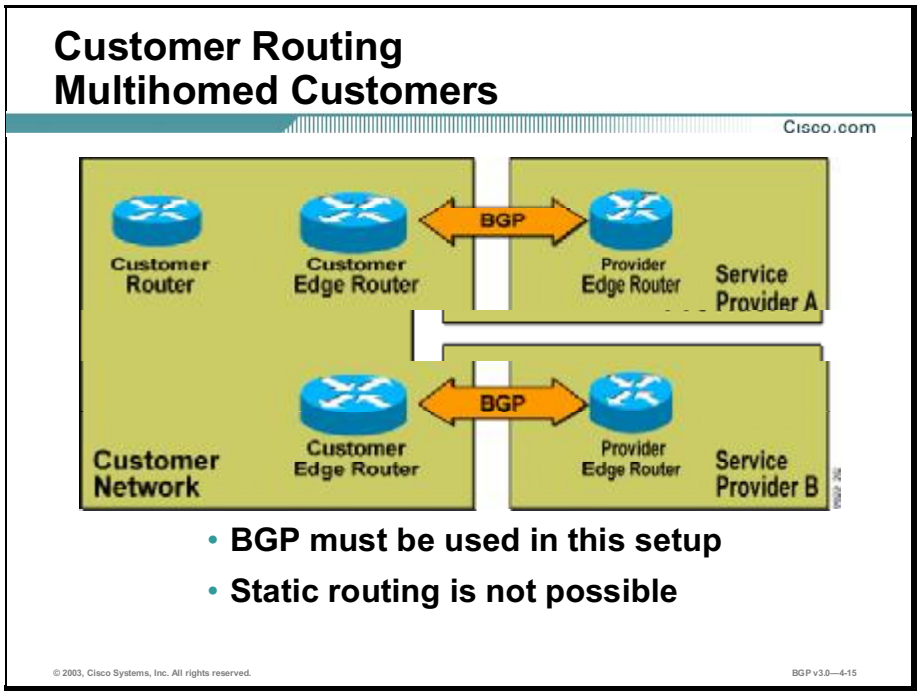
- **You can still use static routing if link and remote equipment failure can be detected reliably**
- **BGP between the customer and the service provider is usually used in this setup**

© 2003, Cisco Systems, Inc. All rights reserved.

You can also use static routing for multiple permanent connections between two different routers on the customer network to two different routers on the service provider network if the failures can be detected by the link-level procedures. When one of the connections is lost, the link-level detects this loss and places the interface in a down state. Because the interface is in the down state, the static route, which points out of the down interface, becomes invalid. As a result, the router stops the redistribution of the static route into BGP.

However, customers requiring the use of multiple connections and multiple routers very often do not rely on the link-level procedures. These customers require a routing protocol such as BGP to detect the failures. Because BGP uses handshaking and reliable transfer, it always detects a failed link or failed remote router.

 Copyright © 2003, Cisco Systems, Inc.



Multiple permanent connections to more than one ISP always require the use of dynamic routing with BGP. The customers requiring this type of connection do not just want to protect the network connectivity from link failures or remote router failures, they also want to protect their network connectivity from serious problems in the network of an ISP.

Monitoring the link status cannot detect a problem inside one of the ISP networks. If the link is still up and the ISP edge router is still up, the link-level procedures do not indicate any problems. However, the ISP network may suffer from severe problems. An ISP network can be partitioned or disconnected from the rest of the Internet without having any problems with the edge router and the access line to the customer network.

The only way to detect this situation is to use BGP with both ISPs and receive full Internet routing from both of them. When one of the ISPs has problems, the edge router, being the BGP neighbor of the customer, withdraws those routes that it can no longer reach. This action means that the customer routers know which Internet routes each ISP can reach at the moment.

Practice

Q1) What will happen if a link failure is not detected where multiple permanent connections between a single router on the customer side and a single router on the ISP side are configured with static routes?

A) Fast failover will occur.

B) The traffic will encounter a black hole.

C) Customers will get a message that the Internet is down.

D) Nothing happens. All routing continues as usual.

Q2) Multiple permanent connections to more than one ISP always require the use of what?

A) dynamic routing

B) mobile routing

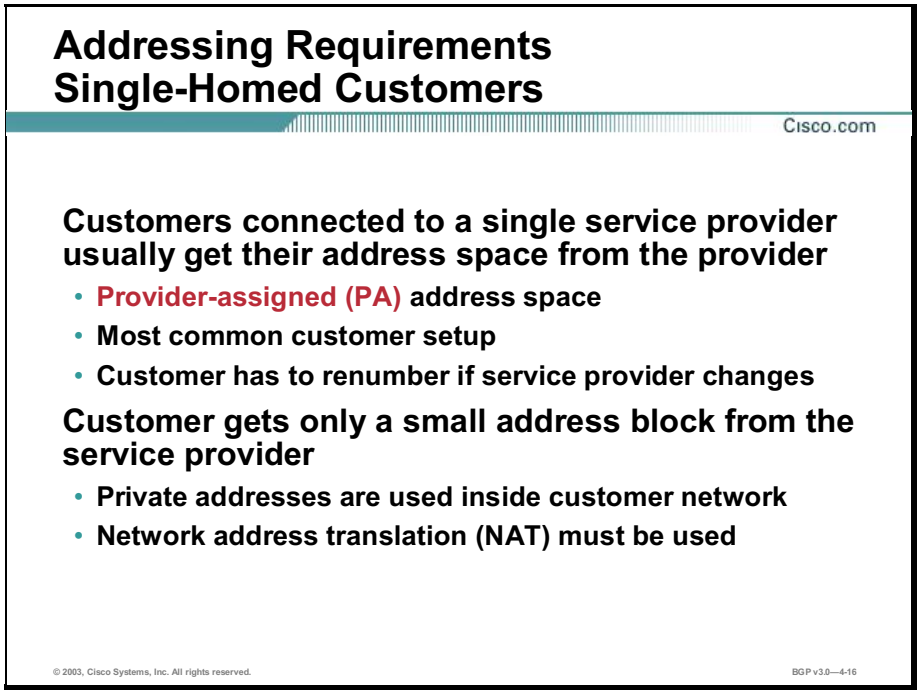
C) static routing

D) secure routing

 Copyright © 2003, Cisco Systems, Inc.

Addressing Requirements

This topic describes the different addressing schemes that customers use to connect to a service provider.

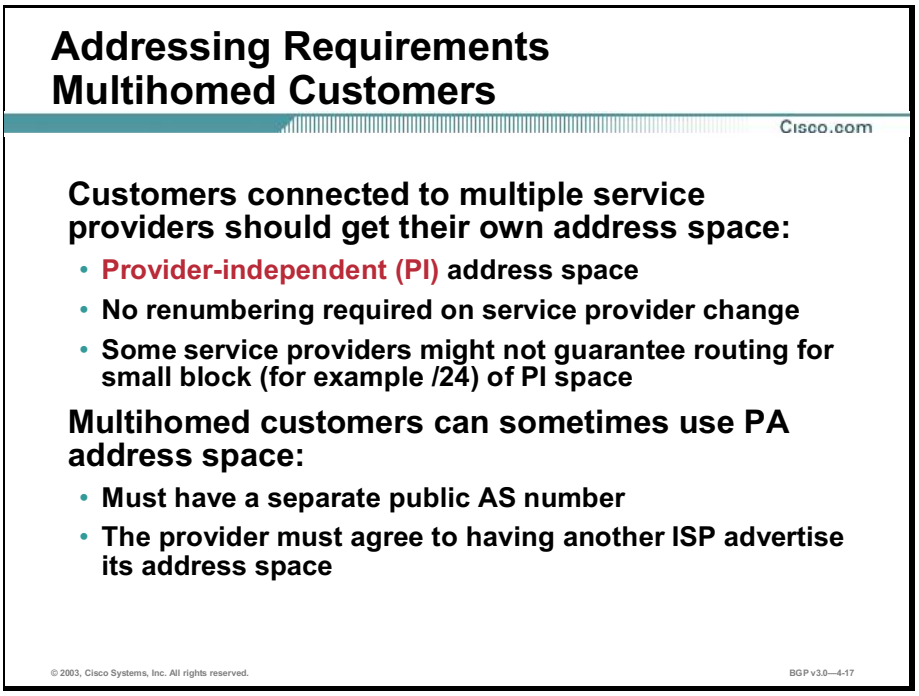

Addressing Requirements
Single-Homed Customers
Cisco.com
Customers connected to a single service provider usually get their address space from the provider
• Provider-assigned (PA) address space
• Most common customer setup
• Customer has to renumber if service provider changes
Customer gets only a small address block from the service provider
• Private addresses are used inside customer network
• Network address translation (NAT) must be used
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-16

Customers connected to a single ISP usually get their address space assigned by the ISP. An ISP is usually assigned a large address space to delegate to its customers. Because all customers of one ISP get their addresses from one or a few address spaces, it is very likely that the ISP is able to aggregate the customer addresses before sending the routes to the rest of the Internet.

Most customers are connected to a single ISP, which means that they are using provider-assigned (PA) addresses. If the customer should decide to change its service provider, the customer must return its PA addresses to the old ISP and receive a new assignment of PA addresses from the new ISP. Otherwise, the ISPs are no longer able to perform efficient address aggregation.

The consequence for the customer is that the customer has to renumber its network when it changes its service provider.

Some customers decide to use private addresses within their network and do Network Address Translation (NAT) at the connection point to the ISP. This setup means that customers require only a very small portion of public addresses assigned by the ISP. In addition to conserving address space for the benefit of the Internet as a whole, this setup also means that when the customer decides to change its service provider, addresses are renumbered only at the NAT point. The rest of the customer network does not need to be renumbered.


Addressing Requirements
Multihomed Customers
Cisco.com
Customers connected to multiple service providers should get their own address space:
• Provider-independent (PI) address space
• No renumbering required on service provider change
• Some service providers might not guarantee routing for small block (for example /24) of PI space
Multihomed customers can sometimes use PA address space:
• Must have a separate public AS number
• The provider must agree to having another ISP advertise its address space
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-17

Customers connected to more than one ISP should, if possible, assign their own address space and not have addresses delegated from any of their ISPs. Such assigned addresses are called provider-independent (PI).

A customer using PI addresses can change its service provider without renumbering its network. The address space is not in any way bound to a particular provider. This arrangement means that no ISP can aggregate the customer routes before sending them to the rest of the Internet. The routes propagate through the Internet with the prefix lengths given.

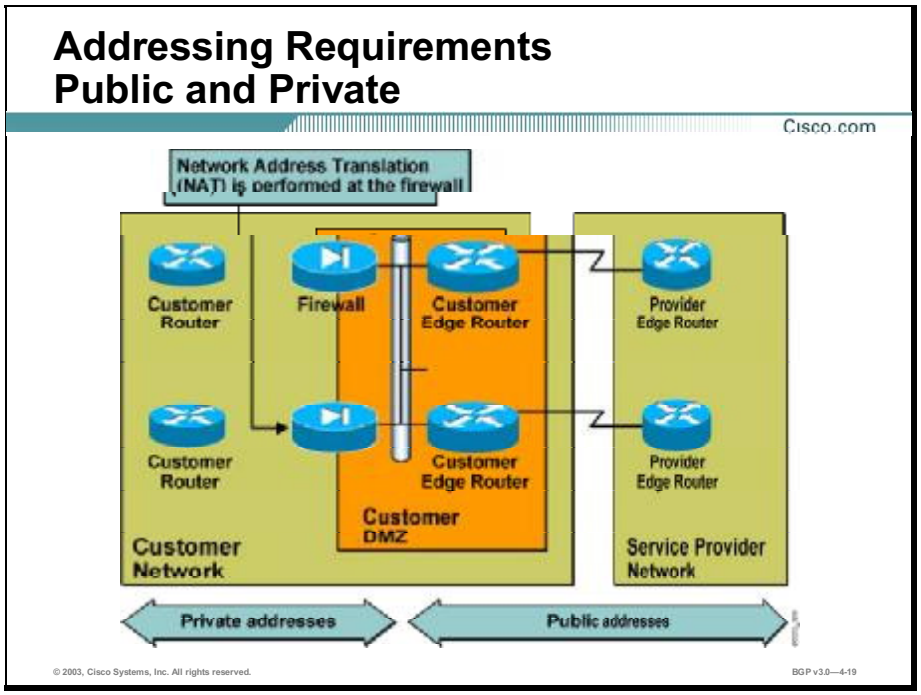
Some large ISPs filter out routes with long prefixes. ISPs do not want to populate their routing tables with a large number of explicit routes that should have been aggregated into a route summary before they were sent to them. As a result, the customer announcing small blocks of PI addresses, which cannot be aggregated, may not be reachable from all parts of the Internet. A larger block of PI addresses solves the problem.

A multihomed customer can in some cases use PA addresses. The address space must be assigned from one of the ISPs. When the customer announces the block of PA addresses to both ISPs, both should propagate the addresses to the rest of the Internet. The provider that assigned the address space should also announce the larger block of addresses, of which the customer is announcing a subset.

Other ISPs now receive two alternate explicit routes and an overlapping route summary. Filtering out explicit routes is more likely at this time because the other ISPs recognize these as routes that can be aggregated. If the other ISPs filter out the more explicit routes, the customer is still reachable as long as both providers are announcing the overlapping route summary.

 Copyright © 2003, Cisco Systems, Inc.

Example



In this example, the customer uses private addresses inside its own network. Only a very small network segment, called the customer demilitarized zone (DMZ), has been assigned public addresses.

The customer network is connected to the customer DMZ using two alternate firewalls with both firewalls doing NAT. All packets leaving the customer network have their addresses translated to a public address belonging to the DMZ subnet. The reverse translation is made in the reverse traffic direction.

In this case, the customer requires only a very small block of public addresses. These addresses can be PA addresses. If the customer decides to change its service provider, renumbering is not a problem because only a few devices need to be reconfigured by the customer.

Care must be taken so that traffic flows symmetrically through the firewalls. Otherwise NAT does not work. The easiest way to achieve this symmetry is to let only one firewall be active at a time.

Practice

Q1) When a customer decides to use private addresses within its network, what must it do to connect to its ISP?

A) No action is required.

B) The customer must summarize its address space.

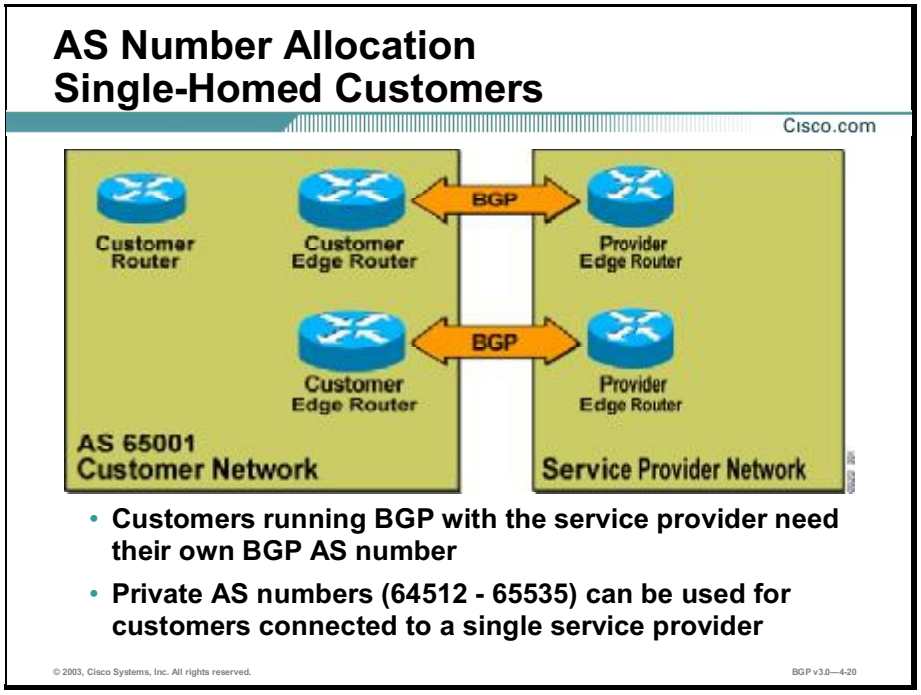
C) The customer must configure static addressing.

D) The customer must use NAT.

 Copyright © 2003, Cisco Systems, Inc.

AS Number Allocation

This topic describes different AS numbering schemes that customers use to connect to a service provider.

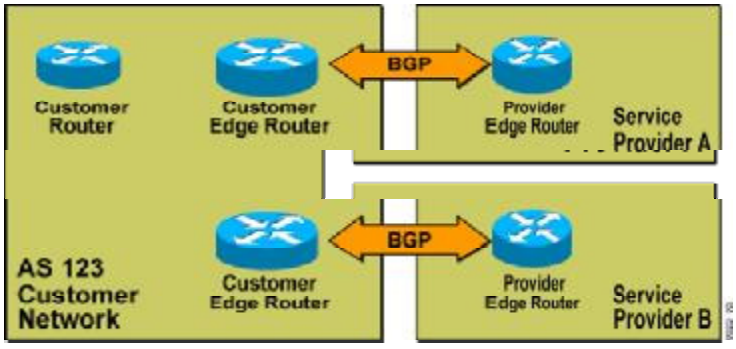

AS Number Allocation
Single-Homed Customers
Cisco.com
Customer Router
Customer Edge Router
BGP
Provider Edge Router
Customer Edge Router
BGP
Provider Edge Router
AS 65001
Customer Network
Service Provider Network
• Customers running BGP with the service provider need their own BGP AS number
• Private AS numbers (64512 - 65535) can be used for customers connected to a single service provider
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-20

BGP requires the use of AS numbers. When BGP is configured, the AS number is mandatory information. However, public AS numbers are a scarce resource. Customers should use public AS numbers only when required. A customer who uses BGP to exchange routing information with only one ISP does not require a public AS number. This customer can use a private AS number.

An ISP network, which is running BGP with some of its customers, must determine whether a public or a private AS number is required for each customer. When the customer can use a private AS number, the ISP must allocate one from the range of private AS numbers (64512 – 65535). The ISP must make sure not to assign any of the private AS numbers to more than one customer.

When the ISP receives BGP routes from the customer, the ISP routers see the private AS number in the AS path and treat the private number as any other AS number. However, before the ISP propagates any of these routes to the rest of the Internet, it must remove the private AS numbers from the AS path, because the same AS number may be in use by someone else. After the private AS number is removed, the route appears as belonging to the public AS of the ISP.

AS Number Allocation Multihomed Customers



- **Multihomed customers must run BGP with their service providers**
- **Multihomed customers must use public AS numbers for their autonomous system**

© 2003, Cisco Systems, Inc. All rights reserved.

A multihomed customer requires a public AS number and must run BGP with both of its ISPs. The customer should not use a private AS number because both ISPs must propagate the customer routes to the rest of the Internet. If the customer does use a private AS number, and both ISPs remove the number before sending it to the rest of the Internet, then the customer routes will appear to be local in the public AS of both ISPs. To make BGP work correctly, multihomed customers need to avoid this situation.

Note	With the help of the AS number translation feature, private AS numbers can also be used for multihomed customers, but this type of configuration is not encouraged.

Multihomed customers are correctly connected to the Internet by assigning a public AS number to the customer network. This public AS appears in the AS path and should be propagated by the service provider to the rest of the Internet. The customer network is now reachable by the rest of the Internet through both providers. The route with the shortest AS path is used by Internet endpoints as the best route to the customer network.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What number ranges correctly indicate private AS numbers?

A) 1024 – 2048

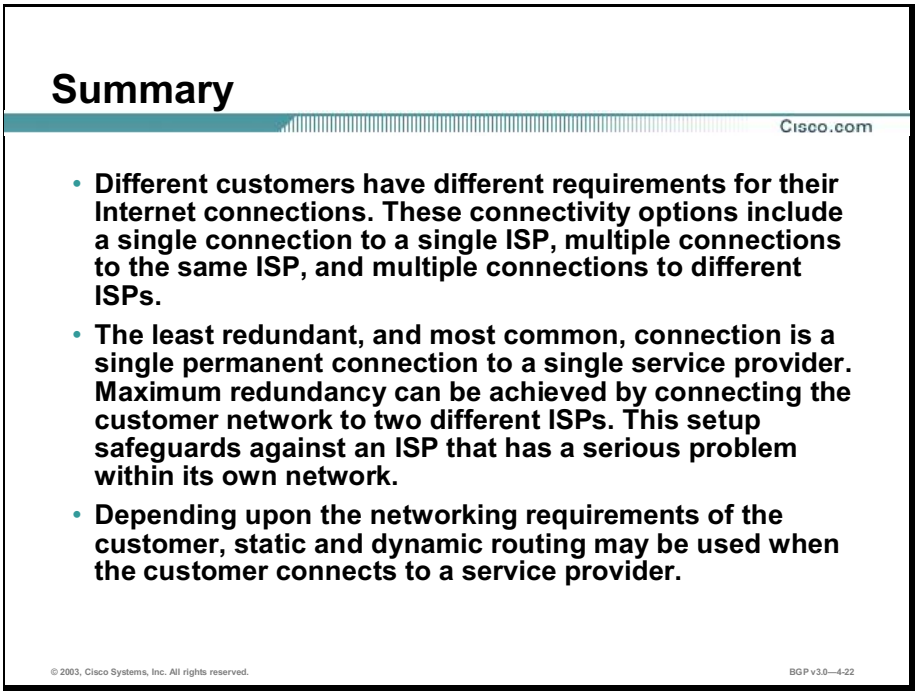
B) 32768 – 64511

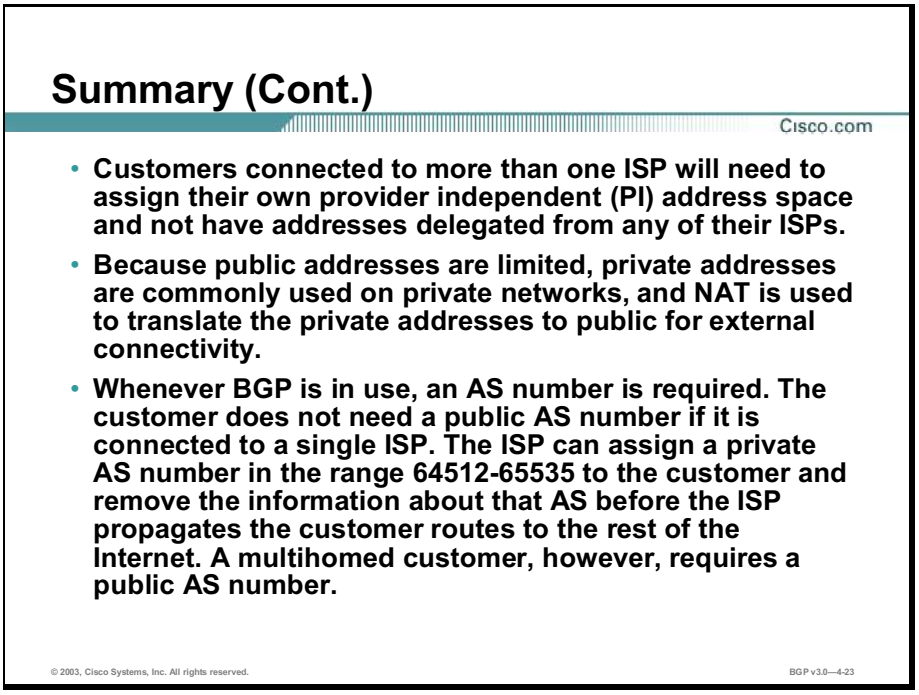
C) 64512 – 65535

D) 65536 – 131072

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Different customers have different requirements for their Internet connections. These connectivity options include a single connection to a single ISP, multiple connections to the same ISP, and multiple connections to different ISPs.
• The least redundant, and most common, connection is a single permanent connection to a single service provider. Maximum redundancy can be achieved by connecting the customer network to two different ISPs. This setup safeguards against an ISP that has a serious problem within its own network.
• Depending upon the networking requirements of the customer, static and dynamic routing may be used when the customer connects to a service provider.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-22


Summary (Cont.)
Cisco.com
• Customers connected to more than one ISP will need to assign their own provider independent (PI) address space and not have addresses delegated from any of their ISPs.
• Because public addresses are limited, private addresses are commonly used on private networks, and NAT is used to translate the private addresses to public for external connectivity.
• Whenever BGP is in use, an AS number is required. The customer does not need a public AS number if it is connected to a single ISP. The ISP can assign a private AS number in the range 64512-65535 to the customer and remove the information about that AS before the ISP propagates the customer routes to the rest of the Internet. A multihomed customer, however, requires a public AS number.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-23

 Copyright © 2003, Cisco Systems, Inc.

Next Steps

After completing this lesson, go to:

- Implementing Customer Connectivity Using Static Routing lesson

References

For additional information, refer to these resources:

- For more information on connecting to a service provider, refer to “The Easy Guide to Selecting an Internet Service Provider” at the following URL: http://www.cisco.com/warp/public/cc/so/cuso/smso/crn/ezgd_pl.htm

- For more information on how NAT works, refer to “How NAT Works” at the following URL: http://www.cisco.com/en/US/tech/tk648/tk361/technologies_tech_note09186a0080094831.shtml

- For more information on configuring connections to service providers, refer to “Sample Configurations for Load Sharing with BGP in Single and Multihomed Environments” at the following URL: http://www.cisco.com/warp/public/459/40.html

Quiz: Customer-to-Provider Connectivity Requirements

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Identify different physical connections that customers use to connect to a service provider
- Describe the levels of redundancy provided by each physical connection type that customers use to connect to a service provider
- Identify different routing schemes that customers use to connect to a service provider
- Describe routing schemes that are appropriate for each physical connection type that customers use to connect to a service provider
- Describe the addressing schemes that customers use to connect to a service provider
- Describe AS numbering schemes that customers use to connect to a service provider

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) If a customer required additional bandwidth as well as redundancy, what method would be preferred?

A) a single permanent connection to one ISP

B) permanent connections to more than one ISP

C) dial-up connections to more than one ISP

D) multiple permanent connections to one ISP

 Copyright © 2003, Cisco Systems, Inc.

Q2) What type of redundancy do multiple permanent connections providing load-sharing configuration display?

A) link

B) equipment

C) service provider

D) routing protocol

Q3) In a customer-to-provider routing scheme, what method of routing is preferred due to its lower complexity?

A) policy-based routing

B) dynamic routing

C) content routing

D) static routing

Q4) Why is it that with multiple permanent connections to more than one ISP, the use of dynamic routing with BGP is required?

A) When one of the connections is lost, the link level detects this loss and places the interface in a down state.

B) Monitoring the link status cannot detect a problem inside one of the ISP networks.

C) Static routes detect problems inside one of the ISP networks.

D) It is not required, and static routing may be used.

Q5) What can be done when a customer is assigned only a very small subnet of public addresses?

A) Purchase more addresses as required.

B) Use NAT.

C) Add a service provider.

D) Add links to the same service provider.

Copyright © 2003, Cisco Systems, Inc.

Q6) What are two different addressing schemes that customers use to connect to a service provider? (Choose two.)

A) provider-independent

B) customer-independent

C) provider-assigned

D) customer-assigned

Q7) Which two of the following criteria are required for a customer to be multihomed to multiple ISPs? (Choose two.)

A) A public AS number.

B) A private AS number.

C) The customer must run BGP with both of its ISPs.

D) The customer must run BGP with one ISP and may use static routing with the other.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Implementing Customer Connectivity Using Static Routing

Overview

Customers connect to the Internet using service providers to enable different applications such as intranet connectivity with Virtual Private Networks (VPNs), extranet connectivity with suppliers, and other Internet applications. Different customers have different connectivity requirements depending upon their business model, redundancy requirements, and even network budget.

This lesson discusses static routing as a solution for connecting customer networks to service providers. Included in this lesson is a discussion of when static routing should be used between a customer and a provider, how to configure static routing in nonredundant, backup, and load-sharing configurations.

Importance

When a single connection to a service provider or multiple connections to the same service provider are the options that a customer has of connecting to the Internet, static routing is the best routing approach to implement between customer and provider. When implementing customer-to-provider connectivity with static routes, it is important that network administrators understand existing guidelines, as discussed in this lesson. Knowledge of static routing implementation guidelines will aid in successfully deploying static routing network configurations.

Objectives

Upon completing this lesson, you will be able to:

- Identify when to use static routing between a customer and a service provider in a BGP environment
- Describe the characteristics of static routing between a customer and a service provider in a BGP environment
- Identify design considerations for propagating static routes in a service provider network
- Configure static route propagation in a BGP environment, given a scenario with different service levels
- Configure a typical backup setup using static routing between a customer and a service provider in a BGP environment
- Describe the limitations of floating static routes when used in typical backup static routing scenarios and the corrective actions to overcome these limitations
- Describe the characteristics of load sharing when you are configuring static routing between a customer and a service provider

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

 Copyright © 2003, Cisco Systems, Inc.

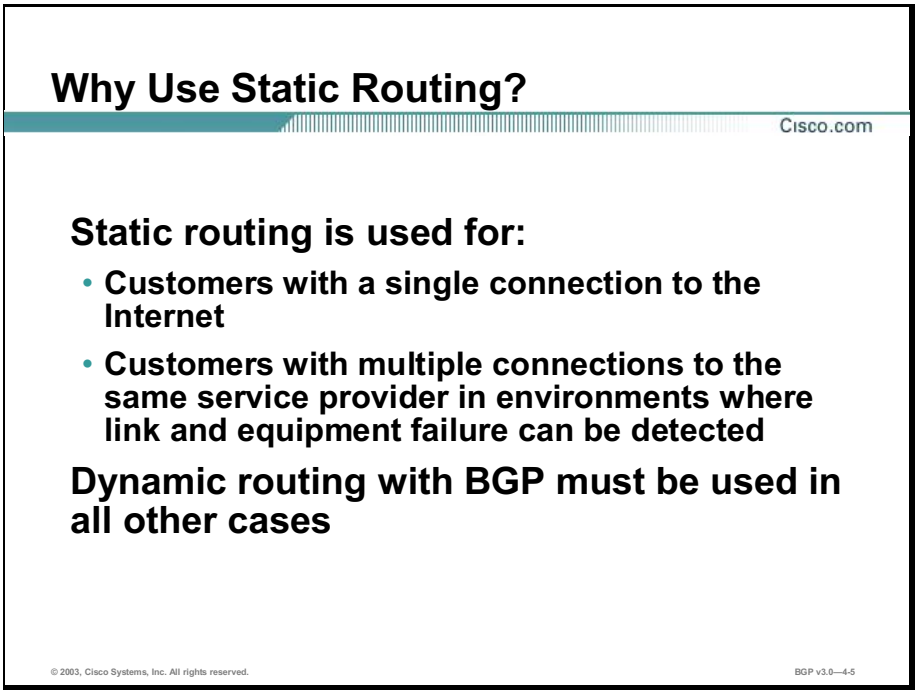
Outline

This lesson includes these topics:

- Overview
- Why Use Static Routing?
- Characteristics of Static Routing
- Designing Static Route Propagation in a Service Provider Network
- Static Route Propagation Case Study Parameters
- BGP Backup with Static Routes
- Floating Static Routes with BGP
- Load Sharing with Static Routes
- Summary
- Assessment (Quiz): Implementing Customer Connectivity Using Static Routing

Why Use Static Routing?

This topic identifies when to use static routing between a customer and a service provider in a Border Gateway Protocol (BGP) environment.


Why Use Static Routing?
Cisco.com
Static routing is used for:
• Customers with a single connection to the Internet
• Customers with multiple connections to the same service provider in environments where link and equipment failure can be detected
Dynamic routing with BGP must be used in all other cases
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-5

Static routing is the best solution to implement when there is no redundancy in the network topology. A single connection between the customer network and the service provider network does not provide any redundancy. If the link goes down, the connection is lost regardless of what routing protocol is configured in the customer or provider network. When there are redundant connections between the customer network and the network of a single service provider, static routing can be used under specific circumstances.

A static, default route must be conditionally announced by the customer edge routers using an Interior Gateway Protocol (IGP). If the link to one of the customer edge routers goes down, then the router must be able to detect the failure and invalidate the static default route. Announcement of this router as a default gateway using an IGP must now cease. Likewise, on the service provider edge routers, the static routes pointing to the customer networks must be invalidated if the link between them goes down, and redistribution to BGP is therefore stopped.

If link-level procedures cannot detect a link failure, the interface remains in the up state. The static routes are not invalidated, and packets are forwarded into a black hole. In this case, since the router cannot detect a failure at the link-level, BGP must be used between the customer and the provider.

BGP must also be used between the customer and the service provider networks when the customer is multihomed. This is the case regardless of what link failure detection mechanisms are in use.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Which of these two situations identify when to use static routing between a customer and a service provider in a BGP environment? (Choose two.)

A) customers with a single connection to the Internet

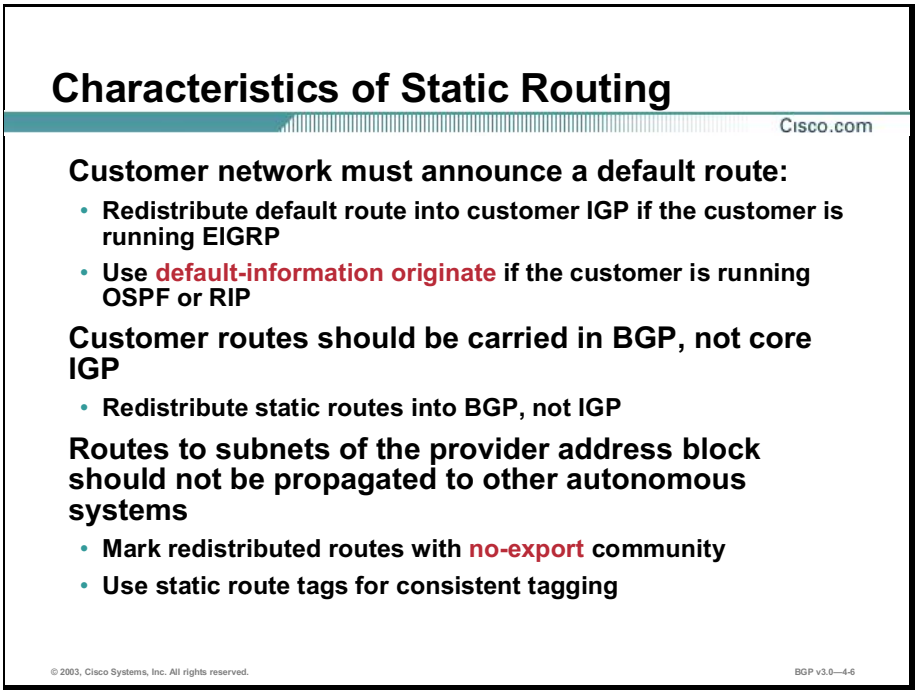
B) customers with multiple connections to multiple service providers

C) customers with multiple connections to the same service provider

D) customers with a single connection to multiple service providers

Characteristics of Static Routing

This topic describes the characteristics of static routing between a customer and a service provider in a BGP environment.


Characteristics of Static Routing
Cisco.com
Customer network must announce a default route:
• Redistribute default route into customer IGP if the customer is running EIGRP
• Use default-information originate if the customer is running OSPF or RIP
Customer routes should be carried in BGP, not core IGP
• Redistribute static routes into BGP, not IGP
Routes to subnets of the provider address block should not be propagated to other autonomous systems
• Mark redistributed routes with no-export community
• Use static route tags for consistent tagging
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-6

When static routing is implemented between the customer network and the Internet service provider (ISP) network, the edge router of the customer must announce itself as a default gateway or a gateway of last resort. This procedure must be done using the IGP within the customer network because different routers within the customer network must be able to select the best route to the exit point of the network.

Different IGPs use different methods of announcing a router as a gateway of last resort. Enhanced Interior Gateway Routing Protocol (EIGRP) uses the concept of default network while Open Shortest Path First (OSPF) and Routing Information Protocol (RIP) send reachability information about network 0.0.0.0/0. In either case, the network operators of the customer network are responsible for configuring their network to use the customer edge router as a gateway of last resort.

When static routing is used between the customer and the provider, the edge router of the provider must propagate a static route that points to the customer network, to all other routers within the ISP network, and also to the rest of the Internet. The network operators in the ISP network propagate the route using a configuration command to start redistributing the routes into BGP.

Customer routes should not be redistributed into the IGP of the ISP network. Care should be taken that the IGP of the ISP network does not carry too many routes. Redistributing customer routes into the IGP could potentially cause poor performance and may eventually cause a complete shutdown of IGP routing at the service provider.

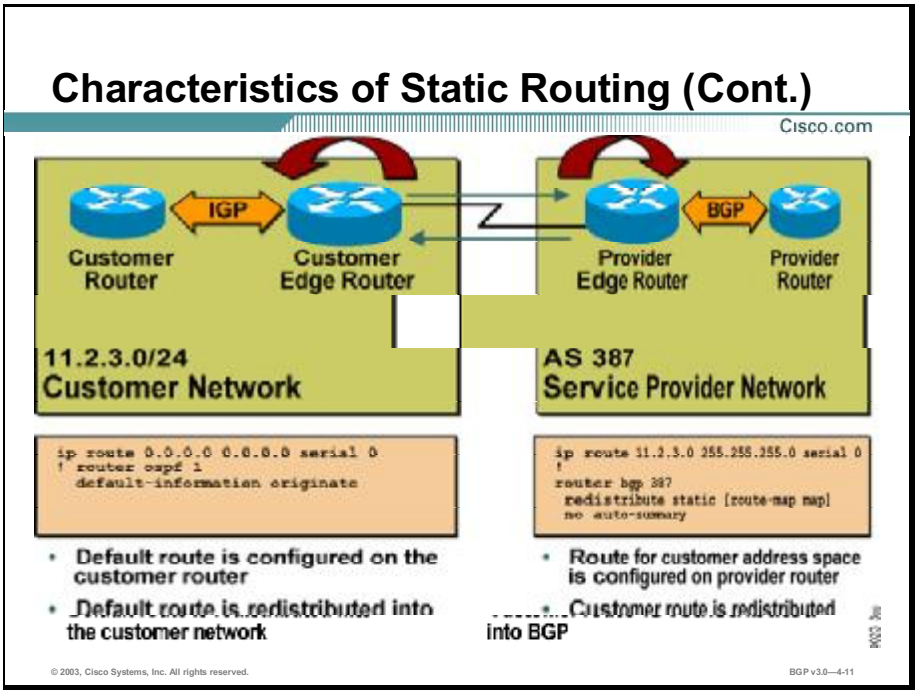
 Copyright © 2003, Cisco Systems, Inc.

If a customer uses provider-assigned (PA) addresses and the ISP announces a large block of addresses where the network of this customer is only a small portion of the block, then the routes of this customer should not be propagated by the service provider to the rest of the Internet. Instead, the rest of the Internet should receive only an announcement containing the larger block of addresses.

An easy way of achieving this setup is to use *communities* within the ISP network. Any customer route that should not be announced to the rest of the Internet is marked using the **no-export** community. To ensure that the BGP communities get propagated, at least over all Internal Border Gateway Protocol (IBGP) sessions, the network operators of the ISP network must configure a **send-community** option for all IBGP neighbors. The edge routers of the ISP network then see the **no-export** community and filter those routes out before sending the update to External Border Gateway Protocol (EBGP) neighbors.

Communities are set using route-maps. A route-map can select routes based on various attributes. One of these attributes is the route tag. Through configuration, a route tag can be assigned by the router to specific static routes. This option means that the network operators of the ISP network can invent a scheme of tagging where all static routes that should not be propagated to other autonomous systems are assigned a specific tag. Then a route-map can select all routes with that tag and assign them the **no-export** community.

Example



In the figure, the customer network is connected to the Internet using a single permanent connection to a single service provider. In this case, a routing protocol does not add any redundancy and would only add complexity.

The customer edge router has a static default route pointing to the interface serial 0. If the serial interface goes down, the route is invalid. The **default-information originate** command is configured in the OSPF process on the customer router; therefore, the router announces a default route into OSPF only as long as it has a valid default route itself.

The service provider edge router also has a static route, declaring the customer IP network number as reachable over the serial 0 interface. It is also invalid if the interface goes into the down state. The ISP edge router must forward this information to all other ISP routers and to the rest of the Internet. This action is accomplished by redistributing the static route into BGP. As long as the static route is valid, BGP announces it. To the rest of the Internet, the customer network appears as reachable within the autonomous system (AS) of the ISP. As far as the rest of the Internet is concerned, the customer is a part of the service provider AS.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What could potentially cause poor performance of the service provider IGP routing?

A) when static routing is used

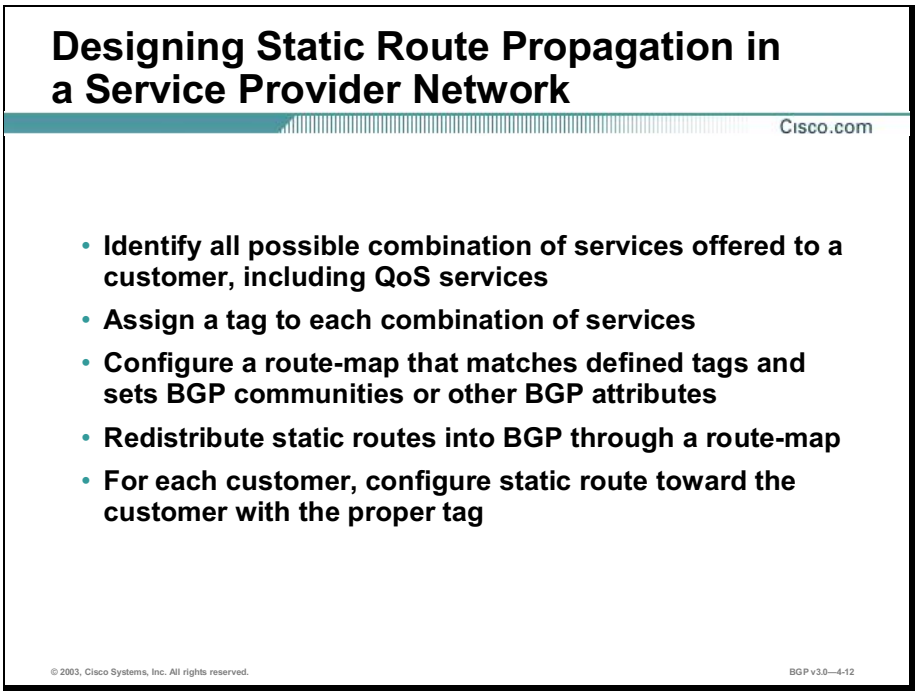
B) redistribution of customer routes into the IGP

C) if a customer uses PA addresses.

D) if the edge router of the customer announces itself as a default gateway

Designing Static Route Propagation in a Service Provider Network

This topic identifies the design considerations for propagating static routes in a service provider network.


Designing Static Route Propagation in a Service Provider Network
Cisco.com
• Identify all possible combination of services offered to a customer, including QoS services
• Assign a tag to each combination of services
• Configure a route-map that matches defined tags and sets BGP communities or other BGP attributes
• Redistribute static routes into BGP through a route-map
• For each customer, configure static route toward the customer with the proper tag
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-12

You can easily extend the principle of using tags when configuring static routes, and assigning different communities based on those tags, to implement a more complex routing policy:

- Identify all different service levels offered to customers and then all the different combinations of these service levels.
- You must assign each combination its own tag value and its own community.
- You must configure a route-map, which selects routes with each and every one of the assigned tags and sets the corresponding community value. Because the processing of a route-map stops when the match clause of a statement is met, each route should only be assigned a single combination of communities. Therefore, you must take great care to assign a tag and a community combination to each combination of services provided.
- When the provider edge routers redistribute static routes into BGP, these routes must pass through the route-map. BGP assigns the correct community depending on the tag values given on the configuration line for each of the static routes.

 Copyright © 2003, Cisco Systems, Inc.

- Finally, you need to configure static routes. Before you configure a static route for a specific customer, you must identify the combination of the services provided to this customer. Then you must look up the corresponding tag value. After you have configured the route, you must assign the tag.

With this routing policy, every static route to a customer network is assigned a tag and the redistributed BGP route is assigned a corresponding community. The BGP communities attached to the routes signal to other routers in the ISP network which particular service combination you should use.

Practice

Q1) What must you identify before you configure a static route for a specific customer?

A) the static routes

B) the corresponding tag value

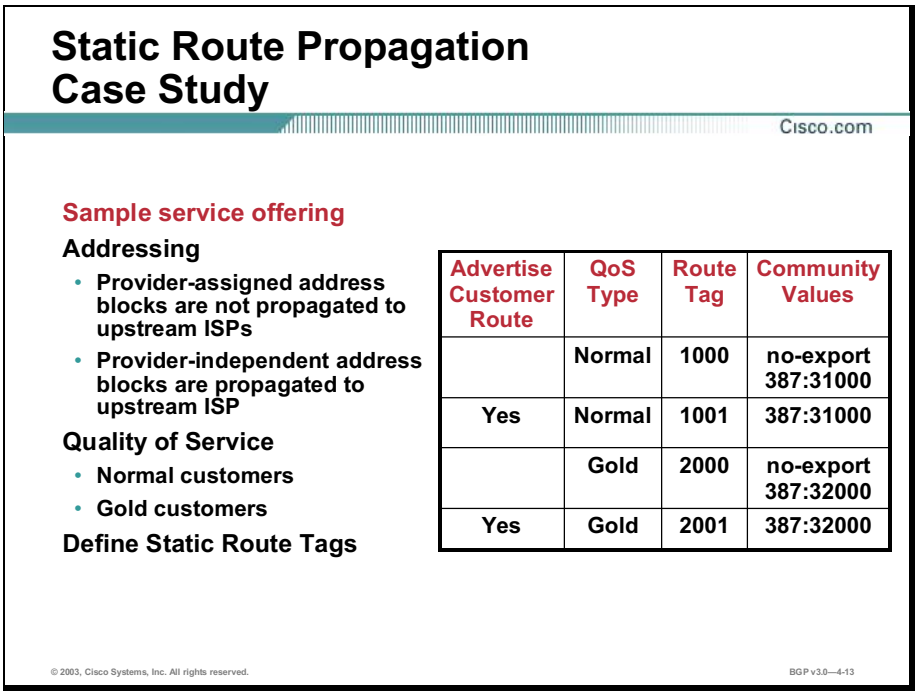
C) the corresponding community value

D) the combination of the services provided to this customer

Copyright © 2003, Cisco Systems, Inc.

Static Route Propagation Case Study Parameters

This topic provides a scenario with varied service levels in which static route propagation is configured in a BGP environment.


Static Route Propagation Case Study
Cisco.com
Sample service offering
Addressing
• Provider-assigned address blocks are not propagated to upstream ISPs
• Provider-independent address blocks are propagated to upstream ISP
Quality of Service
• Normal customers
• Gold customers
Define Static Route Tags
Advertise Customer Route | QoS Type | Route Tag | Community Values
Normal | 1000 | no-export 387:31000
Yes | Normal | 1001 | 387:31000
Gold | 2000 | no-export 387:32000
Yes | Gold | 2001 | 387:32000
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-13

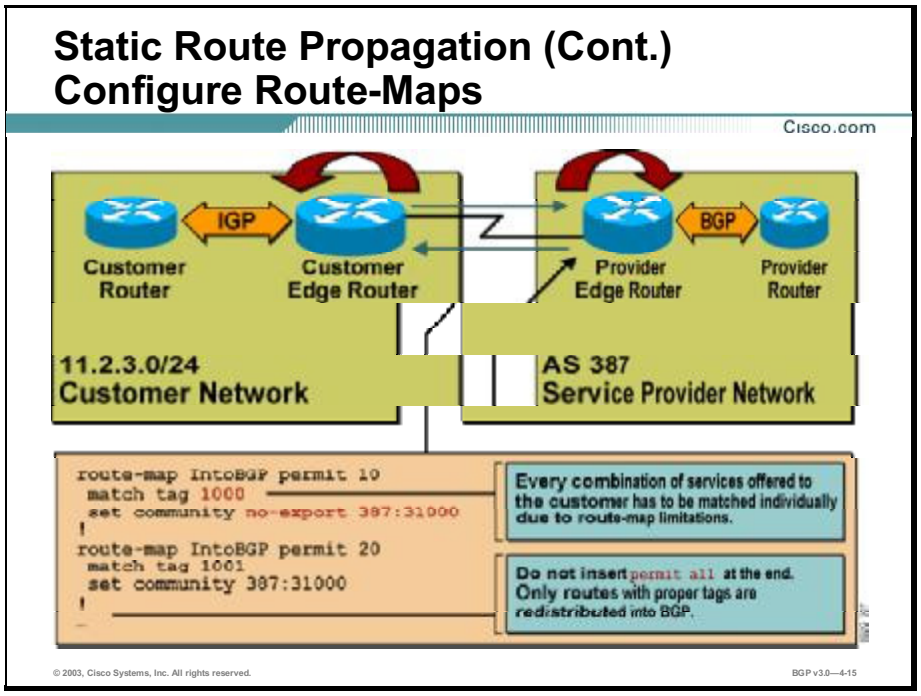
In this case study, the service provider offers two different service levels to its customers: Normal and Gold. Customers are also assigned IP address blocks. Some customers use PA addresses, which the ISP does not announce as explicit routes. The large route-summary block announced by the ISP covers these customers. Other customers use PI addresses that must be explicitly announced to the Internet by the service provider.

Because there are two different quality of service (QoS) services, Normal and Gold, and both PA and PI addresses, the total number of combinations to cover the network policy is four:

- Normal QoS routes, which are assigned by the ISP and should not be explicitly announced
- Normal QoS routes, which are PI routes and should be explicitly announce
- Gold QoS routes, which are assigned by the ISP and should not be explicitly announced
- Gold QoS routes, which are PI and should be explicitly announced

Each of these four combinations receives its own tag value and its own community combination.

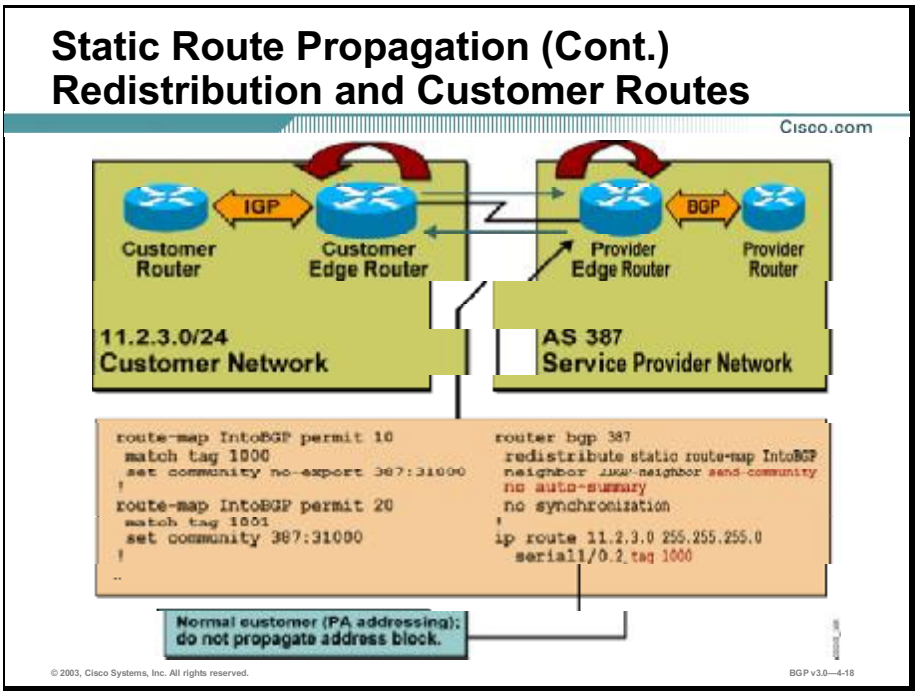
 Copyright © 2003, Cisco Systems, Inc.



Network operators configure a route-map in the ISP edge router that has the static routes to the customer network. Redistribution of the configured customer static routes into BGP is also performed at the ISP edge router.

Because a route-map can match an individual route in one single **route-map** statement only, a single tag value, representing each combination of services, must be assigned to the static routes by the router. When a route is matched, the interpretation of the route-map for that individual route stops. The route-map has one statement for each combination, and each statement matches a tag value and assigns the corresponding community combination for that tag.

The route-map is applied during the redistribution of customer static routes into BGP at the provider edge router. Because the route-map has no **permit any** statement at the end, those static routes not assigned any of the tags being used are not redistributed. The route-map filters these routes out, forcing the network operators to make a tag assignment to all customer routes. Furthermore, the route-map filtering may help catch administrator configuration entry errors, thus giving each and every customer the service combination that they are entitled to.

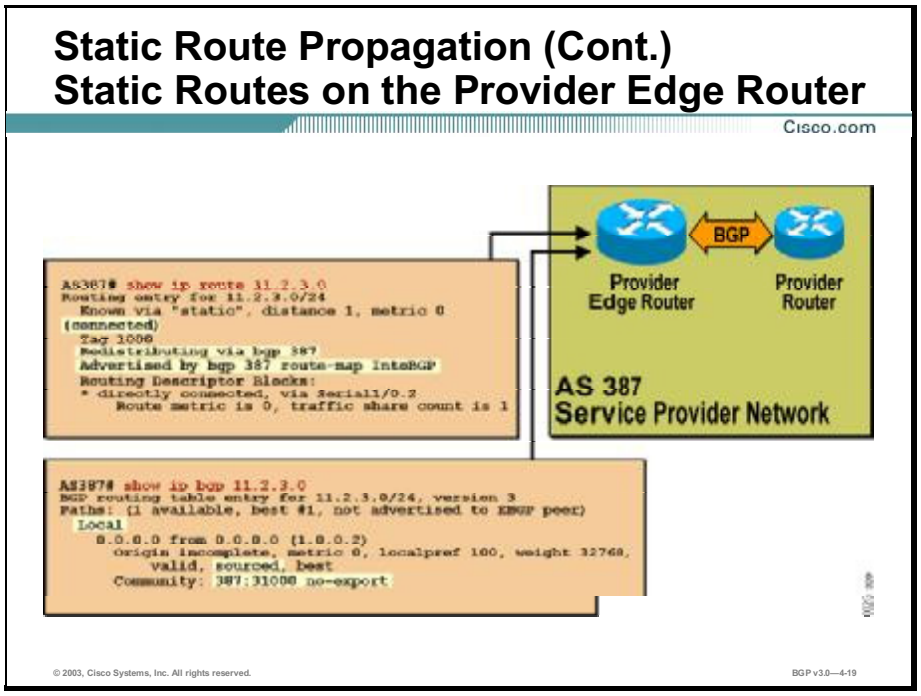


This figure shows how the service provider edge router uses the route-map named *IntoBGP* when redistributing the static routes into BGP. Because the route-map assigns community values that will be used by other routers within the ISP network, network operators must configure all IBGP neighbors with the **send-community** qualifier.

Use the **no auto-summary** BGP configuration command to avoid having the subnet 11.2.3.0/24 automatically summarized into 11.0.0.0/8.

Now, when connecting customers, the network operators identify which service combination to use for this particular customer. The three services associated with this particular customer are: apply normal QoS, use a provider assigned network number, and the provider should not explicitly announce the customer routes. A static route to the customer is configured and assigned the appropriate tag value of 1000, which represents the specified services assigned to the customer.

 Copyright © 2003, Cisco Systems, Inc.


Static Route Propagation (Cont.)
Static Routes on the Provider Edge Router
Cisco.com
AS387# show ip route 11.2.3.0
Routing entry for 11.2.3.0/24
Known via "static", distance 1, metric 0
(connected)
Tag 1000
Redistributing via bgp 387
Advertised by bgp 387 route-map IntoBGP
Routing Descriptor Blocks:
* directly connected, via Serial1/0.2
Route metric is 0, traffic share count is 1
AS387# show ip bgp 11.2.3.0
BGP routing table entry for 11.2.3.0/24, version 3
Paths: (1 available, best #1, not advertised to EBGP peer)
Local
0.0.0.0 from 0.0.0.0 (1.0.0.2)
Origin incomplete, metric 0, localpref 100, weight 32768,
valid, sourced, best
Community: 387:31000 no-export
BGP
Provider Edge Router
Provider Router
AS 387
Service Provider Network
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-19

The **show ip route** command displays information from the forwarding table about subnet 11.2.3.0/24. The route is learned by static configuration and is redistributed via BGP. The router, through the use of a statically assigned tag, has assigned a tag value of 1000 to the customer route, and the route must pass through route-map *IntoBGP* before being inserted into the BGP table.

The **show ip bgp** command displays information from the BGP table about subnet 11.2.3.0/24. The route is local within this AS and is sourced by this router. The BGP communities 387:31000 and **no-export** have been assigned by the router to the redistributed customer route using the provider defined route-map prior to inserting the customer route into the BGP table.

Practice

Q1) How many tag value(s) must you assign to each combination of services?

A) 1

B) 2

C) 16

D) 32

Q2) What do you use at the edge router of a service provider to redistribute static routes into BGP?

A) **no-export**

B) **send-community**

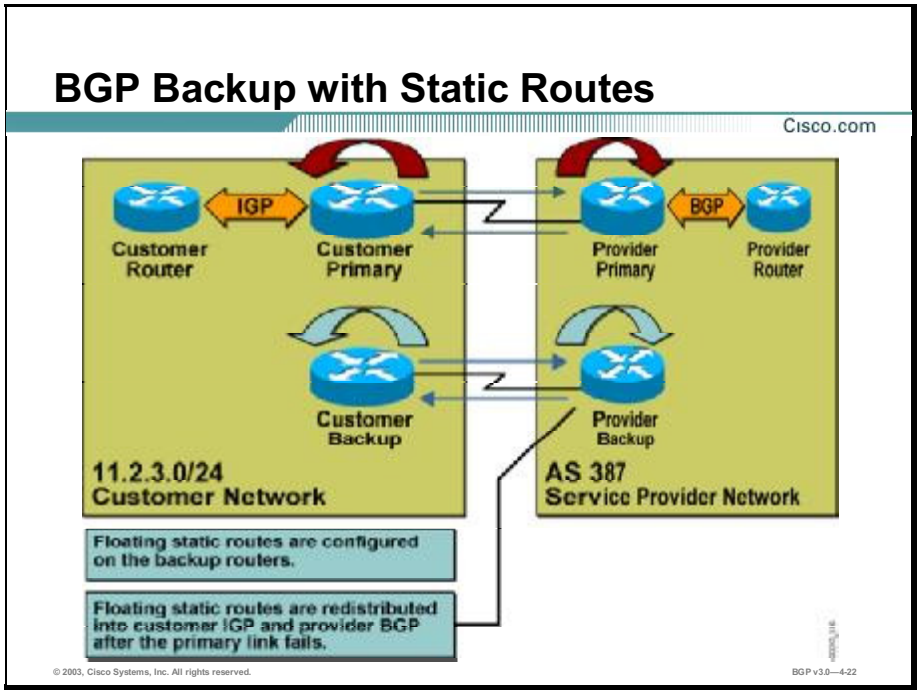
C) **route-map**

D) **route-community**

 Copyright © 2003, Cisco Systems, Inc.

BGP Backup with Static Routes

This topic explains how to configure a typical backup setup using static routing between a customer and a service provider in a BGP environment.



This example illustrates a case where the customer network has two connections to a single service provider. One connection between the customer network and the ISP is the primary connection, and the other connection is used for backup purposes only. If link-level procedures can detect link failures and a failure in the remote router, then static routing can be used instead of a dynamic routing protocol between the customer and provider networks.

As in the previous example, where no backup link was available, the primary edge router of the customer has a static default route toward the ISP and the primary edge router of the ISP has static routes toward the customer. The customer router redistributes the static default route into its IGP. The ISP router redistributes the static routes into BGP.

If the primary link goes down, the link-level procedures set the interface to the down state, causing the static routes pointing out through the interface to be invalid, and removing the routes from the forwarding table. When the interface changes back to the up state, the static route will reappear in the forwarding table.

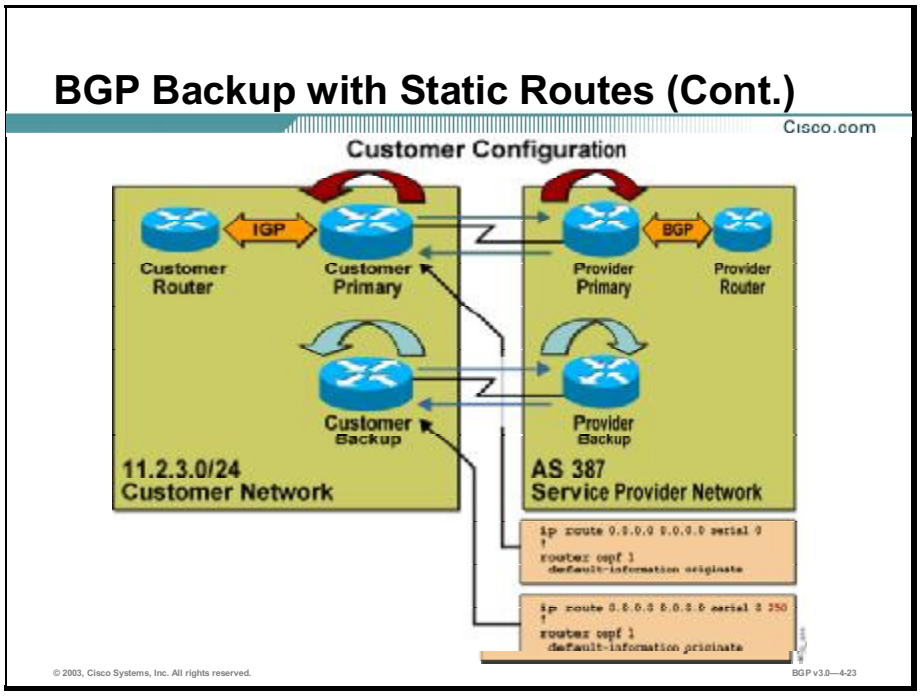
Redistribution of routes into any routing protocol is conditioned by the appearance of the route in the forwarding table. Thus, if the interface goes down, the router removes the static route from its forwarding table, and the route is withdrawn from the routing protocol. When the static route reappears, the redistribution process inserts it into the routing protocol again.

The backup edge router of the customer also uses a default static route toward the ISP, via the backup link. The backup edge router is also redistributing the default route into the IGP. However, the static route used is a floating static route, which is assigned a high administrative

distance (AD), higher than the AD of the customer IGP. As long as the primary link works, the IGP provides the customer backup edge router with the primary default route. Due to a higher AD, the backup static default route is not installed into the backup router forwarding table. Because the static route is not in the forwarding table, it is not redistributed. If the primary link fails, the IGP no longer feeds the backup edge router with a default route. The backup static default route is the only remaining default route. Therefore, the router will install the floating default route into its forwarding table, and subsequently redistribute it into the IGP.

The backup edge router of the ISP can also use floating static routes, which are redistributed into the ISP BGP process.

 Copyright © 2003, Cisco Systems, Inc.

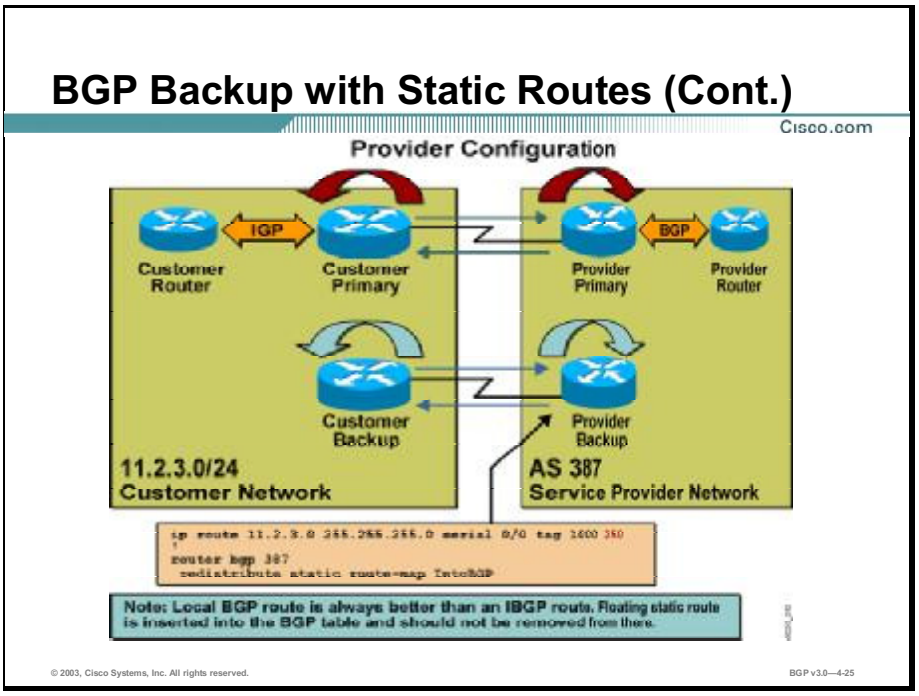


In the figure, the customer network and the ISP network are connected using leased lines with High-Level Data Link Control (HDLC) encapsulation. Both the primary and the backup edge routers in the customer network have a static default route toward the serial interface leading to the ISP. Both routers also do redistribution of the default route into the Open Shortest Path First (OSPF) protocol, which is being used as an IGP within the customer network.

However, the static default route in the backup edge router is configured with an AD value set to 250. This AD value is higher than the AD values of any routing protocol. This configuration means that as long as the backup router receives the default route by OSPF, the static default route is not used.

When the primary link goes down, the static default route in the primary router is not valid. The OSPF protocol stops announcing the default route, because the **default-information originate** command is like a redistribution of the default route from the forwarding table of the router.

The backup router now installs its static default route in the forwarding table. The conditions for announcing the default route by OSPF are met and the rest of the customer routers see the backup router as the gateway of last resort.



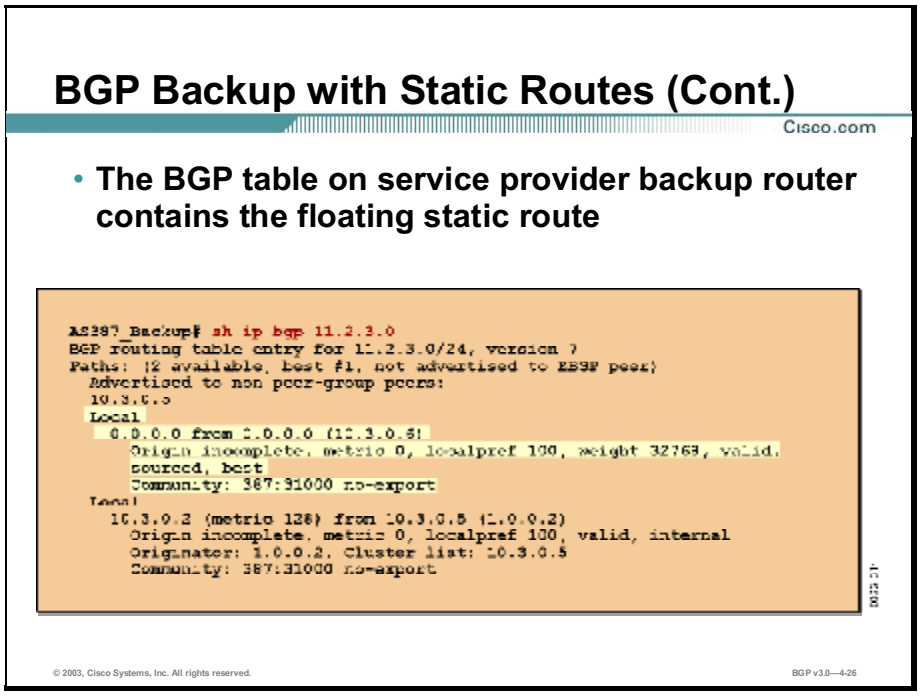
When floating static routes are configured on the provider edge routers, they are also redistributed into BGP. This configuration makes things a little bit more complicated.

The network operator configures a floating static route to the customer subnet 11.2.3.0/24. In the provider edge router, the floating static route is assigned the same tag value as the tag value being used in the primary router. The route-map *IntoBGP* is the same as in the primary router and provides the routes to the customer network with the same communities (the same QoS level and indication whether to explicitly announce it to the rest of the Internet).

The floating static route is configured with an AD value set to 250. This value is higher than any routing protocol. When the backup edge router of the ISP no longer receives any routing protocol information about the customer networks, the router will automatically install the floating static route and subsequently redistribute it into BGP.

Based on BGP route selection rules, the redistributed floating static route will always remain the preferred path if additional BGP configuration is not performed on the provider edge router. This preference means that regardless of whether the primary link comes back, the backup router selects the locally sourced route as the best route. Therefore, the backup router continues to announce a path toward the customer network. The backup link does not go back to the Idle state.

 Copyright © 2003, Cisco Systems, Inc.


BGP Backup with Static Routes (Cont.)
Cisco.com
• The BGP table on service provider backup router contains the floating static route
AS387_Backup# sh ip bgp 11.2.3.0
BGP routing table entry for 11.2.3.0/24, version 7
Paths: (2 available, best #1, not advertised to EBGP peer)
Advertised to non peer-group peers:
10.3.0.5
Local
0.0.0.0 from 0.0.0.0 (11.2.3.5)
Origin incomplete, metric 0, localpref 100, weight 32768, valid,
sourced, best
Community: 387:31000 no-export
Local
10.3.0.2 (metric 128) from 10.3.0.5 (1.0.0.2)
Origin incomplete, metric 0, localpref 100, valid, internal
Originator: 1.0.0.2, Cluster list: 10.3.0.5
Community: 387:31000 no-export
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-26

In this example, the **show ip bgp** command is used in the backup edge router of the provider to display the information about the customer network 11.2.3.0/24. The primary link has come back, so the backup router now sees two alternate routes. The first route is the route that the router itself has redistributed into BGP using the floating static route. This route is locally sourced by this AS and has been assigned a weight value of 32768. The second route is the one received by IBGP from the primary edge router. This AS also sources this route but no weight value is assigned.

The BGP route selection algorithm selects the route with weight value 32768 as the best. As a result, the route received from the primary edge router is not a candidate to be installed in the forwarding table and never competes with the floating static route. The floating static route stays in the forwarding table, and redistribution of the route continues until the backup link goes down and the route is invalid.

Practice

Q1) What is configured on backup routers and redistributed into the customer IGP and provider BGP after a primary link fails?

A) weighted routes

B) floating static routes

C) floating dynamic routes

D) dynamic forwarding table

Q2) When you are configuring a backup router with BGP static routes, how can you ensure that as long as the backup router is receiving the default route from the IGP that the BGP static route is not used?

A) Assign the static route a low AD value.

B) Assign the static route a high AD value.

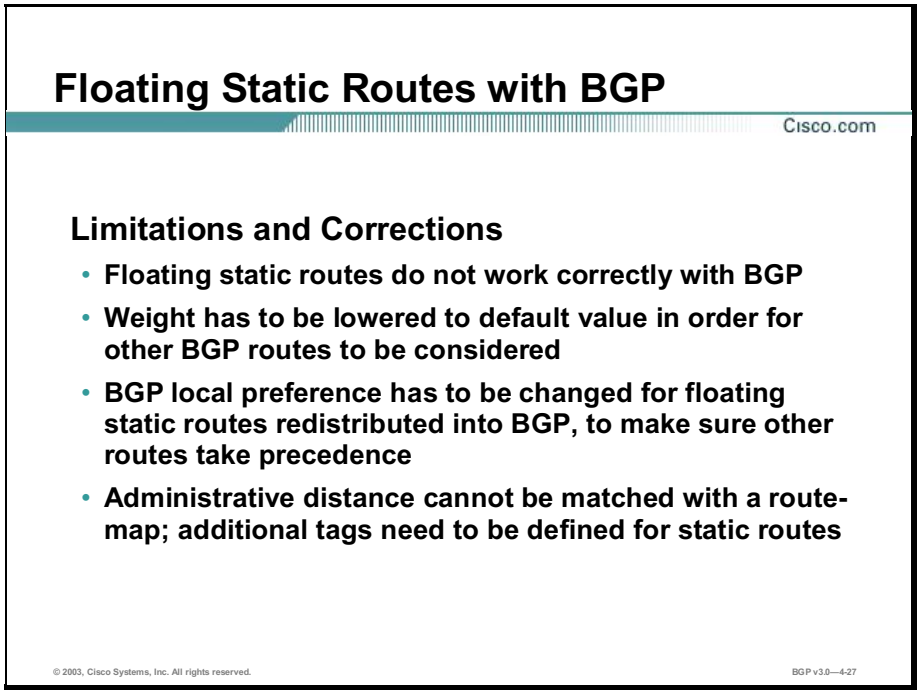
C) Assign the static route an AD equal to that of the IGP.

D) Nothing, the route selection process is automatic.

 Copyright © 2003, Cisco Systems, Inc.

Floating Static Routes with BGP

This topic describes the limitations of floating static routes when used in typical backup static routing scenarios and the corrective actions to overcome these limitations.


Floating Static Routes with BGP
Cisco.com
Limitations and Corrections
• Floating static routes do not work correctly with BGP
• Weight has to be lowered to default value in order for other BGP routes to be considered
• BGP local preference has to be changed for floating static routes redistributed into BGP, to make sure other routes take precedence
• Administrative distance cannot be matched with a route-map; additional tags need to be defined for static routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-27

Unfortunately, floating static routes do not work correctly with BGP. After they are inserted, the floating static route is never removed from the forwarding table even if the primary link comes back.

Whenever you use floating static routes in combination with redistribution into BGP, you will need to take additional configuration steps ensure that the BGP route selection algorithm selects the primary route as the best BGP route when it reappears:

- When a router redistributes a floating static route into BGP, the weight value assigned to the floating static route must be reduced. Otherwise, the floating static route will always be selected as the best BGP route after the first failure of the primary link occurs.

- Local preference values must be also be assigned by the router to the floating static route, giving it a lower local preference than the primary route. This assignment makes sure that the primary route is selected as the best BGP route after it comes back.

These two requirements must be specified on the provider edge router in the route-map *IntoBGP* used for the redistribution. The route-map must select the floating static routes and set weight and local preference. However, a route-map cannot do matching based on the AD value assigned to a static route. Some other means are required to make it possible for the route-map to distinguish between normal static routes that should have normal weight and local preference and the floating static ones that should have their values modified.

The solution is to create additional tag values for this set of static routes. The tag value must not only reflect the QoS level and whether to announce the route, but also the tag value must indicate if it is a primary route or a backup route.

 Copyright © 2003, Cisco Systems, Inc.

Floating Static Routes with BGP (Cont.)

Cisco.com

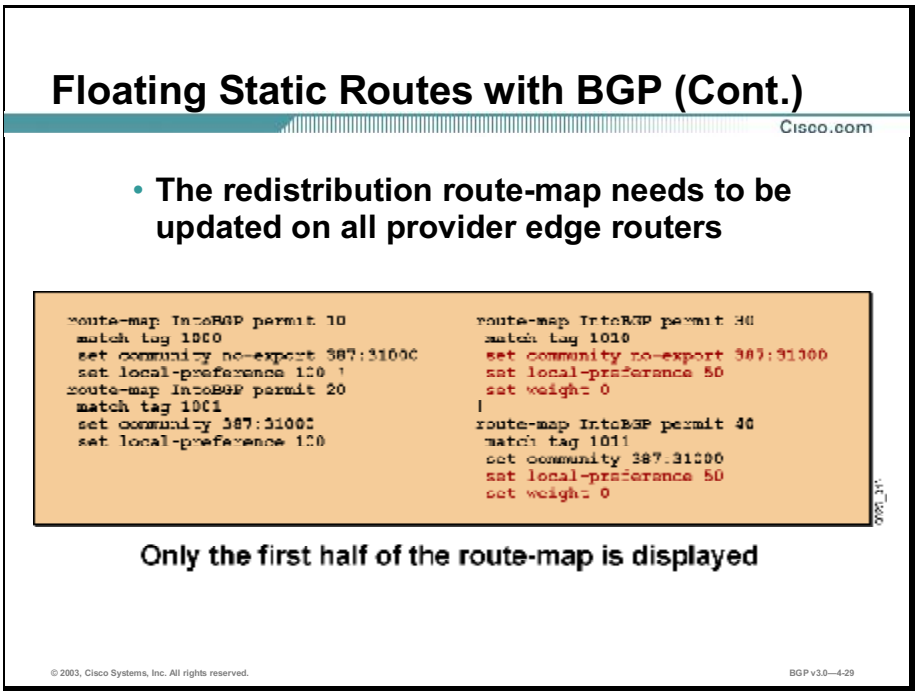
Sample Static Route Tags with Backup

Advertise Customer Route	Backup	QoS Type	Tag	Community Values	Local Preference
		Normal	1000	no-export 387:31000	100
	Yes	Normal	1010	no-export 387:31000	50
Yes		Normal	1001	387:31000	100
Yes	Yes	Normal	1011	387:31000	50
		Gold	2000	no-export 387:32000	100
	Yes	Gold	2010	no-export 387:32000	50
Yes		Gold	2001	387:32000	100
Yes	Yes	Gold	2011	387:32000	50

© 2003, Cisco Systems, Inc. All rights reserved.

There are now eight different tag values identified. Each of tag values indicates a specific combination of explicit route propagation (backup or primary) and QoS level.

When network operators configure static routes in the provider edge router, they must consider which of the combinations that they should use for the route. The route-map that they use when redistributing the static routes into BGP must be configured to recognize all eight combinations and to set the appropriate weight, and community and local preference values.


Floating Static Routes with BGP (Cont.)
Cisco.com
• The redistribution route-map needs to be updated on all provider edge routers
route-map IntoBGP permit 10
match tag 1000
set community no-export 387:31000
set local-preference 120
route-map IntoBGP permit 20
match tag 1001
set community 387:31000
set local-preference 120
route-map IntoBGP permit 40
match tag 1010
set community no-export 387:31000
set local-preference 50
set weight 0
route-map IntoBGP permit 40
match tag 1011
set community 387:31000
set local-preference 50
set weight 0
Only the first half of the route-map is displayed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-29

The configuration output in this figure displays the first half of the route-map *IntoBGP*. The output shows how four of the eight different tags are identified by match clauses. For each of the tag values, the route-map sets the community, the local preference, and, in some cases, the weight.

Because the displayed half of the route-map deals only with those four tags that indicate QoS Normal, all statements in the configuration display have set the BGP community attribute to 387:31000. The part of the route-map, not shown, deals with the four tags that indicate QoS Gold, which would be configured to set the BGP community attribute to 387:32000.

Tag values of 1000, 1010, 2000, and 2010 indicate that the route should not be explicitly propagated. Those routes that should not be explicitly advertised by the provider to the rest of the Internet are assigned the **no-export** community by the route-map.

Tag values 1010, 1011, 2010, and 2011 all indicate that the route is a backup route. Those tags have their weight value set to 0 and their local preference value set to 50. These settings ensure that upon the return of a failed primary route, the provider edge router will select the primary route as its best path and remove the backup, floating static route, from its route table.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Which two of the following statements indicate limitations of floating static routes within BGP? (Choose two.)

A) Weight values cannot be modified.

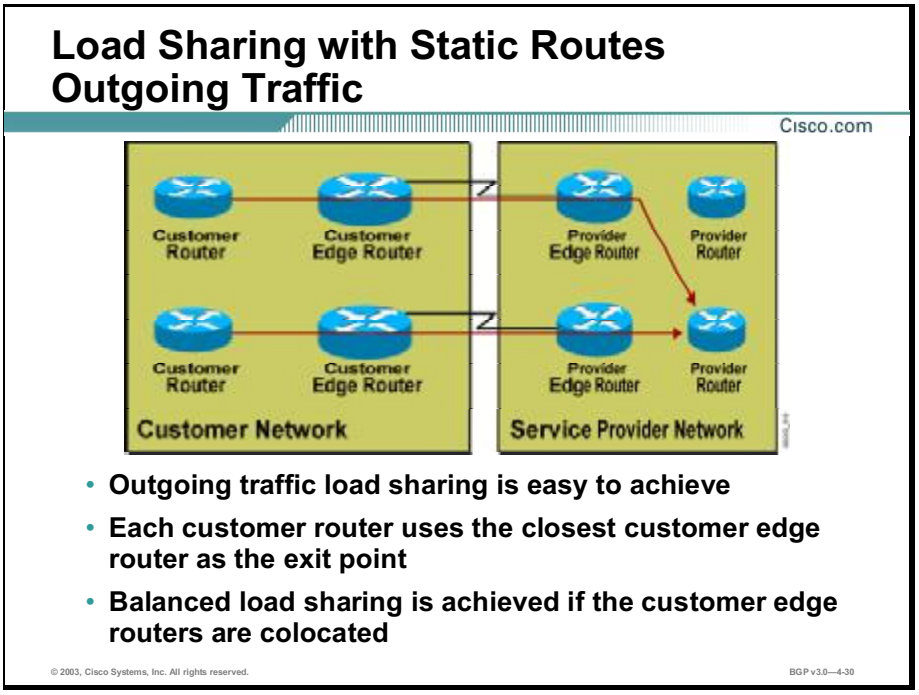
B) Floating routes do not work correctly.

C) AD cannot be matched with a route-map.

D) There is no way to define additional tags for static routes.

Load Sharing with Static Routes

This topic describes the characteristics of load sharing when you are configuring static routing between a customer and a service provider.


Load Sharing with Static Routes
Outgoing Traffic
Cisco.com
Customer Router
Customer Edge Router
Provider Edge Router
Provider Router
Customer Router
Customer Edge Router
Provider Edge Router
Provider Router
Customer Network
Service Provider Network
• Outgoing traffic load sharing is easy to achieve
• Each customer router uses the closest customer edge router as the exit point
• Balanced load sharing is achieved if the customer edge routers are colocated
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-30

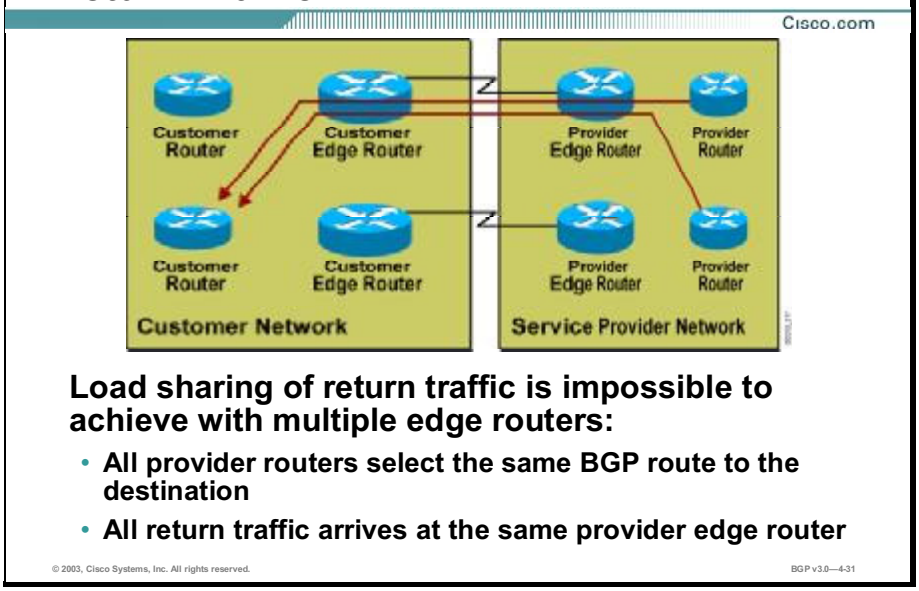
Load sharing of outgoing customer traffic is accomplished by configuring a standard default static route in both customer edge routers. Each static route is valid as long as the serial link in each router is up. When both static routes are valid, both customer edge routers announce the default route into the customer network.

The remaining routers in the customer network see two candidate gateways of last resort. These remaining routers choose the closest one, with respect to the IGP metric. The part of the network that is closer to the uppermost exit point uses that exit point for all outgoing traffic. The other part of the network uses the other (lower) exit point.

If both exit points are colocated, they are equally distant from each of the other routers in the customer network. Each router within the customer network therefore uses load sharing of traffic sent out both exit points.

 Copyright © 2003, Cisco Systems, Inc.

Load Sharing with Static Routes Return Traffic



Load sharing of return traffic is impossible to achieve with multiple edge routers:

- **All provider routers select the same BGP route to the destination**
- **All return traffic arrives at the same provider edge router**

© 2003, Cisco Systems, Inc. All rights reserved.

The provider routers receive routes toward the customer network via BGP. BGP in its default behavior selects a single route as the best route, allowing no load sharing. The provider routers that receive the same BGP route from two edge routers will always select the closer edge router (if all other BGP attributes are equal, the IBGP route with the closer next hop is selected). The part of the ISP network that is closer to the uppermost connection uses that connection. The other part of the ISP network uses the other (lower) connection.

If both connection points are colocated, all provider routers select the same IBGP route based on router-ID (because the IGP metrics are always equal) and all the return traffic is sent over a single link toward the customer network, resulting in no load sharing.

Note	Since Cisco IOS® version 12.2, the IBGP Multipath load-sharing feature enables the BGP-speaking router to select multiple IBGP paths as the best paths to a destination. The best paths or multipaths are then installed in the IP routing table of the router.

Load Sharing with Static Routes Optimizing Return Traffic

Cisco.com

You can optimize return traffic load sharing

- **Each provider edge router advertises only part of customer address space into the provider backbone**
- **Every provider edge router also advertises the whole customer address space for backup purposes**

Load sharing is not optimal—every link will carry return traffic for part of customer address space

© 2003, Cisco Systems, Inc. All rights reserved.

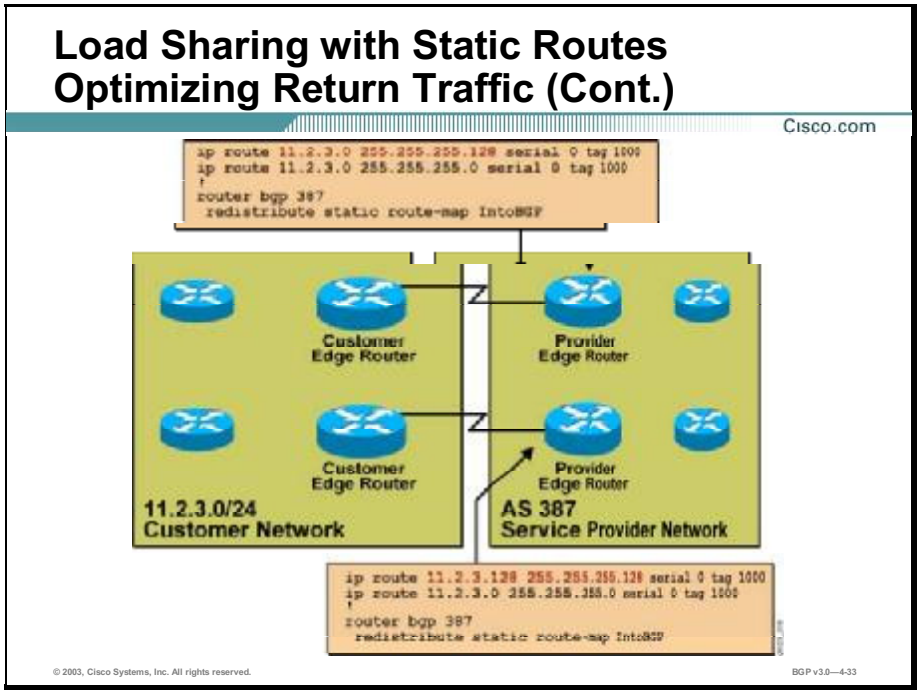
To obtain better control of the return traffic load, the customer address space must be advertised to the provider edge routers using multiple, more explicit routes. The upper edge router could advertise half the address space, and the lower edge router could advertise the other half. For backup reasons, they also should both advertise the entire address space as a larger route summary.

As long as both paths are available, the traffic from the ISP to the customer uses the most explicit route. In this case, two explicit routes are used to send traffic representing one half of the address space over one link and traffic representing the other half of the address space over the other link.

Load sharing in this way does not result in an equal load on the links but rather a statistically based distribution of the traffic load over the links.

 Copyright © 2003, Cisco Systems, Inc.

Example



In the example here, the customer address space 11.2.3.0/24 is partitioned into two smaller blocks: 11.2.3.0/25 and 11.2.3.128/25.

The upper provider edge router advertises the route to 11.2.3.0/25, and the lower router advertises the route to 11.2.3.128/25. Both edge routers also advertise the entire address space 11.2.3.0/24.

The routers in the ISP network direct traffic with destination addresses in the 11.2.3.0/25 range to the upper connection point. Traffic to destinations in the 11.2.3.128/25 range is directed to the lower connection point.

Practice

Q1) When you are performing load sharing of outgoing traffic with static routes, what is the effect of colocating the edge routers?

A) greater throughput

B) faster convergence

C) higher availability

D) balanced load sharing

Q2) When you are using static routes, with what two routing tricks can you optimize return traffic load sharing? (Choose two.)

A) Each provider edge router advertises only part of the customer address space into the provider backbone.

B) Each provider edge router advertises the entire address space of the customer into the provider backbone.

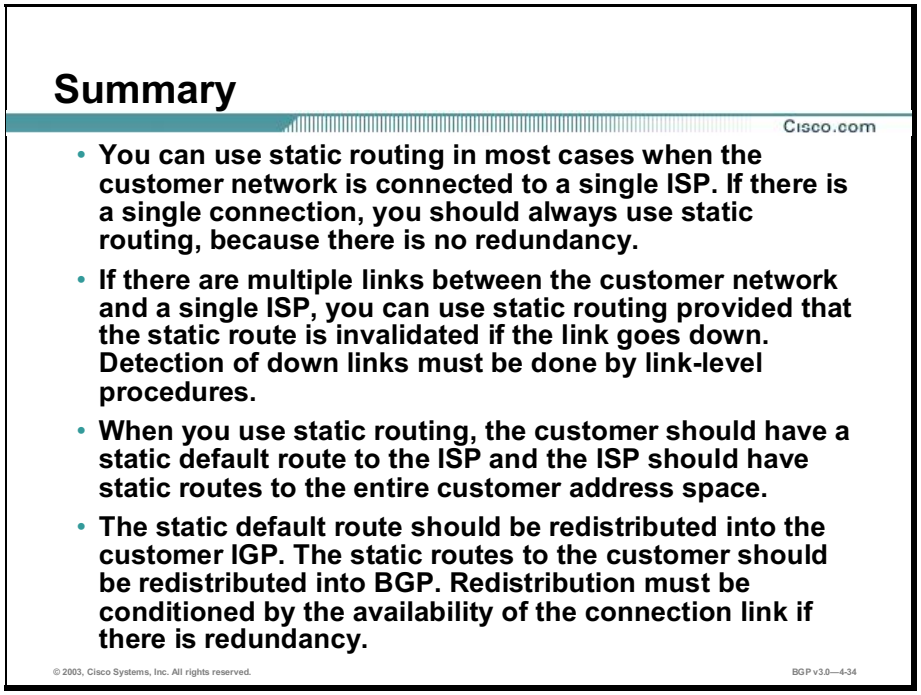
C) Each provider edge router advertises the entire address space of the customer for backup purposes.

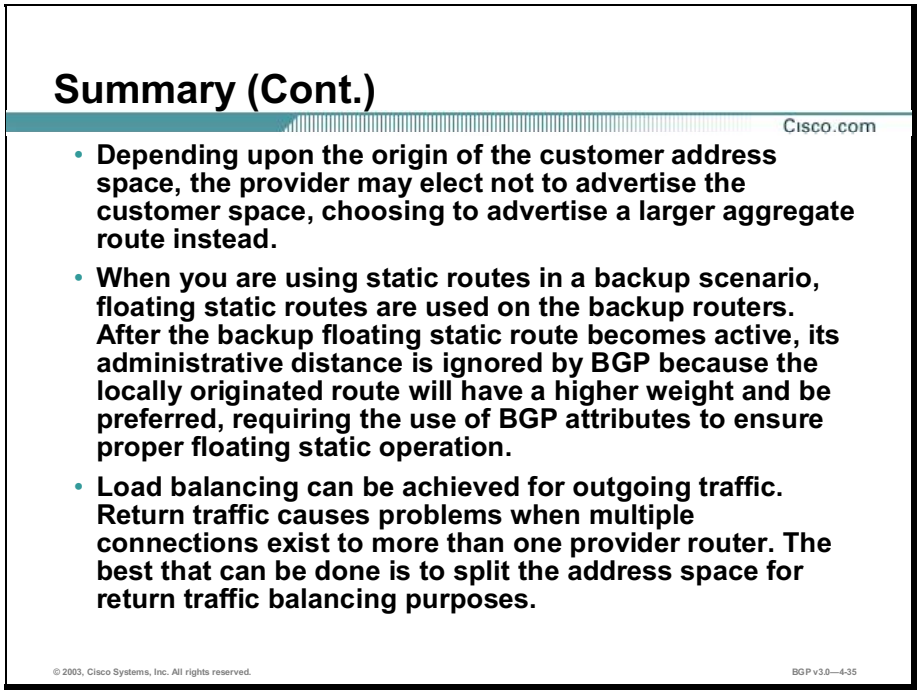
D) Each provider edge router advertises only part of the customer address space for backup purposes.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• You can use static routing in most cases when the customer network is connected to a single ISP. If there is a single connection, you should always use static routing, because there is no redundancy.
• If there are multiple links between the customer network and a single ISP, you can use static routing provided that the static route is invalidated if the link goes down. Detection of down links must be done by link-level procedures.
• When you use static routing, the customer should have a static default route to the ISP and the ISP should have static routes to the entire customer address space.
• The static default route should be redistributed into the customer IGP. The static routes to the customer should be redistributed into BGP. Redistribution must be conditioned by the availability of the connection link if there is redundancy.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-34


Summary (Cont.)
Cisco.com
• Depending upon the origin of the customer address space, the provider may elect not to advertise the customer space, choosing to advertise a larger aggregate route instead.
• When you are using static routes in a backup scenario, floating static routes are used on the backup routers. After the backup floating static route becomes active, its administrative distance is ignored by BGP because the locally originated route will have a higher weight and be preferred, requiring the use of BGP attributes to ensure proper floating static operation.
• Load balancing can be achieved for outgoing traffic. Return traffic causes problems when multiple connections exist to more than one provider router. The best that can be done is to split the address space for return traffic balancing purposes.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-35

Next Steps

After completing this lesson, go to:

- Connecting a Multihomed Customer to a Single Service Provider lesson

References

For additional information, refer to these resources:

- For more information on multihoming, refer to “Sample Configuration for BGP with Two Different Service Providers (Multihoming)” at the following URL: http://www.cisco.com/warp/public/459/27.html

- For more information on static routing, refer to “Routing TCP/IP Volume 1,” Jeff Doyle, Cisco Press, 1998. ISBN 1-57870-041-8

 Copyright © 2003, Cisco Systems, Inc.

Quiz: Implementing Customer Connectivity Using Static Routing

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Identify when to use static routing between a customer and a service provider in a BGP environment
- Describe the characteristics of static routing between a customer and a service provider in a BGP environment
- Identify design considerations for propagating static routes in a service provider network
- Configure static route propagation in a BGP environment, given a scenario with different service levels
- Configure a typical backup setup using static routing between a customer and a service provider in a BGP environment
- Describe the limitations of floating static routes when used in typical backup static routing scenarios and the corrective actions to overcome these limitations
- Describe the characteristics of load sharing when you are configuring static routing between a customer and a service provider

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Copyright © 2003, Cisco Systems, Inc.

Q1) What are two circumstances where you can use static routing as part of installing redundant connections between the customer network and a single service provider network? (Choose two.)

A) The router must be able to detect a link failure.

B) The default route must be announced using the customer IGP.

C) If one link goes down, the interface must remain in an up state.

D) The customer IGP must continue to advertise the static default route.

Q2) A customer route that should not be announced to the rest of the Internet is marked using what?

A) a route tag

B) the **export** community

C) the **no-export** community

D) the public address filter

Q3) When you are designing static route propagation in a service provider network, what three steps must you take? (Choose three.)

A) Assign a tag to each combination of services.

B) Configure a community that matches defined tags.

C) Redistribute static routes into BGP through a route-map.

D) Identify all possible combinations of services offered to a customer.

Q4) What does a route-map assign that will be used by other routers within a network?

A) a tag

B) community values

C) public addressing

D) QoS

 Copyright © 2003, Cisco Systems, Inc.

Q5) What three key pieces of information can you derive from the following router command output? (Choose three.)

```
AS387_Backup# sh ip bgp 11.2.3.0
BGP routing table entry for 11.2.3.0/24, version 7
Paths: (2 available, best #1, not advertised to EBGP peer)
  Advertised to non peer-group peers:
  10.3.0.5
  Local
    0.0.0.0 from 0.0.0.0 (10.3.0.6)
      Origin incomplete, metric 0, localpref 100, weight 32768,
valid,
      sourced, best
      Community: 387:31000 no-export
  Local
    10.3.0.2 (metric 128) from 10.3.0.5 (1.0.0.2)
      Origin incomplete, metric 0, localpref 100, valid, internal
      Originator: 1.0.0.2, Cluster list: 10.3.0.5
      Community: 387:31000 no-export
```

A) The primary link has come back, so the backup router now sees two alternate routes.

B) The primary link has not come back up, but the backup router still sees two alternate routes.

C) The first route is the route that the router itself has redistributed into BGP using the floating static route. This route is locally sourced by this AS and has been assigned a weight value of 32768.

D) The second route is the one received by IBGP from the primary edge router. This AS also sources this route, but no weight value is assigned.

Q6) What two things can you do to overcome the problems that occur when a floating static route is redistributed into BGP? (Choose two.)

A) You must raise the weight value.

B) You must lower the weight value.

C) You must set the AD at a higher value than all other routes.

D) You must assign local preference values, giving the floating static route a lower local preference value than the primary route.

Q7) What are three characteristics of using static routes during load sharing of outgoing traffic? (Choose three.)

A) Outgoing traffic load sharing is easy to achieve.

B) Each customer router uses the closest customer edge router as the exit point.

C) Balanced load sharing is achieved if the customer edge routers are colocated.

D) Local preference values must be assigned, giving the floating static route a lower local preference value than the primary route.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Connecting a Multihomed Customer to a Single Service Provider

Overview

Customers connect to the Internet using service providers to enable different applications such as intranet connectivity with Virtual Private Networks (VPNs), extranet connectivity with suppliers, and other Internet applications. Different customers have different connectivity requirements depending upon their business model, redundancy requirements, and even network budget.

This lesson discusses the use of multiple connections between a customer and a single service provider for backup and load-sharing purposes. Included in this lesson is a discussion of how to configure a customer network and a provider network to accommodate multiple connections between them. Also discussed in this lesson are topics specific to networks with multiple connections between a customer and a single provider such as private autonomous system (AS) number removal and configuration of a network to support either backup links or load sharing (balancing).

Importance

When multiple connections to the same service provider are the only means that a customer has of connecting to the Internet, it is important that the connections are correctly configured to ensure proper interaction between the customer and service provider network. It is also important to understand how to configure routing protocols so that customer backup or load-balancing requirements are met.

Objectives

Upon completing this lesson, you will be able to:

- Configure BGP on a customer network to establish routing between a multihomed customer and a single service provider
- Configure conditional advertising of a customer address space when BGP is used to establish routing between a multihomed customer and a single service provider
- Configure BGP on a service provider network to establish routing between a multihomed customer and a single service provider
- Disable the propagation of private AS numbers to EBGP peers in a service provider network where a multihomed customer is advertising private numbers in the AS path
- Configure a typical backup setup between a multihomed customer and a single service provider in a BGP environment
- Describe how load sharing can be implemented between a multihomed customer and a single service provider in a BGP environment
- Identify the Cisco IOS® command required to configure load sharing between a multihomed customer and a single service provider using BGP multipath
- Configure load sharing between a multihomed customer and a single service provider using EBGP multihop

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

 Copyright © 2003, Cisco Systems, Inc.

Outline

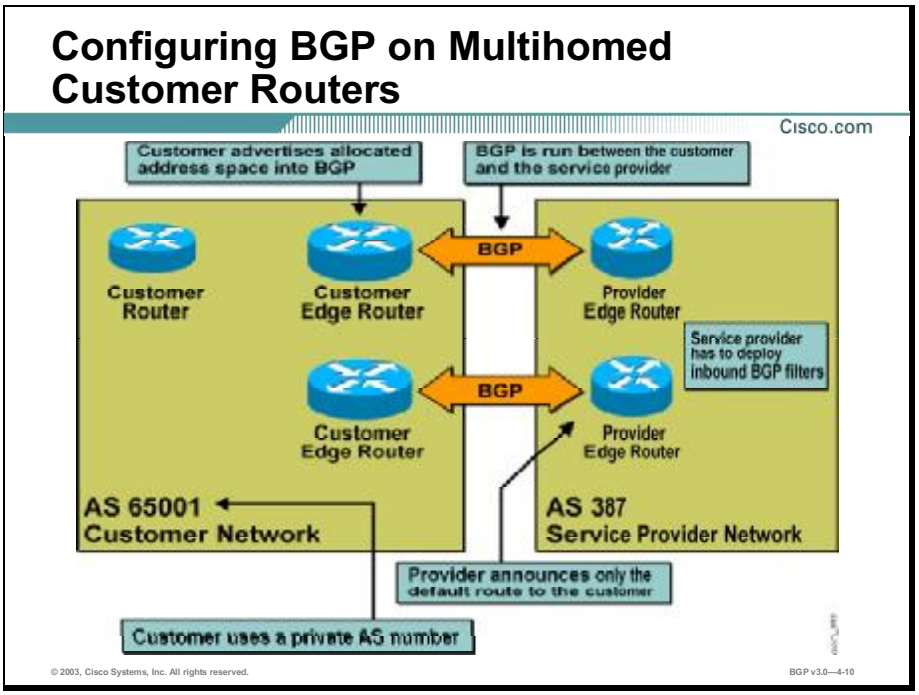
This lesson includes these topics:

- Overview
- Configuring BGP on Multihomed Customer Routers
- Conditional Advertising in Multihomed Customer Networks
- Configuring BGP on Service Provider Routers
- Removing Private AS Numbers
- Backup Solutions with BGP
- Load Sharing with the Multihomed Customer
- Load Sharing with BGP Multipath
- Load Sharing with EBGP Multihop
- Summary
- Assessment (Quiz): Connecting a Multihomed Customer to a Single Service Provider

Copyright © 2003, Cisco Systems, Inc.

Configuring BGP on Multihomed Customer Routers

This topic describes how to configure Border Gateway Protocol (BGP) on a customer network to establish routing between a multihomed customer and a single service provider.


Configuring BGP on Multihomed Customer Routers
Cisco.com
Customer advertises allocated address space into BGP
BGP is run between the customer and the service provider
BGP
Customer Router
Customer Edge Router
Provider Edge Router
Service provider has to deploy inbound BGP filters
BGP
Customer Edge Router
Provider Edge Router
AS 65001 Customer Network
AS 387 Service Provider Network
Provider announces only the default route to the customer
Customer uses a private AS number
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-10

In the case study here, the customer network is connected to a service provider network using multiple permanent links. BGP is used to exchange routing information between the customer and the provider.

Selecting BGP as the routing protocol between the customer and provider network ensures that a link failure or the failure of a remote router is detected. In this scenario, the customer does not require the use of a public AS number or full Internet routing. Instead, a private AS number is assigned to the customer network, and the Internet Service Provider (ISP) sends a default route to the customer through BGP.

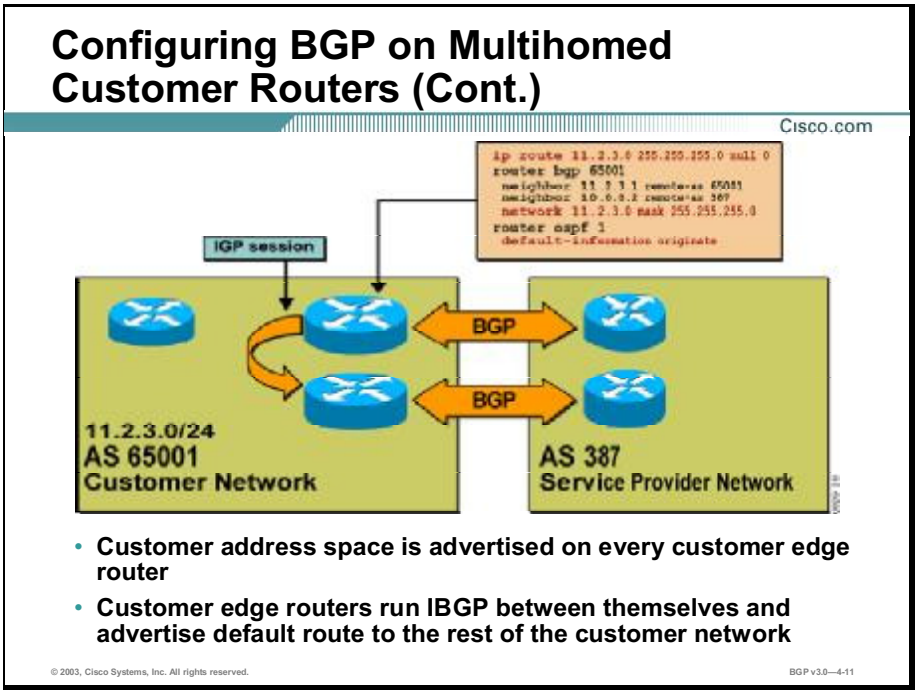
The big difference in this case as compared to a network scenario where static routes and redistribution are used, is that routers within the private AS of the customer now advertise the customer routes via BGP. Thus, the customer is responsible for announcing its own address space. The ISP receives routes from the customer and conditionally propagates them (similar to static routing). If the customer uses provider–assigned (PA) address space, and the ISP can summarize the address space, it will not propagate the explicit routes from the customer to the Internet. The private AS number in the AS-path attribute must first be removed before the ISP can propagate any of the customer routes.

 Copyright © 2003, Cisco Systems, Inc.

Because the customer is now creating BGP routes that are received by the ISP, any error made by the customer can influence routing operation within the ISP network and, if propagated, within the Internet as whole. Announcing a route to a network, which the customer has not been assigned, may induce routing problems. There is always a risk that such routing problems can occur in a service provider network. However, the risk is much greater when the customer, whose network administrators usually have less experience with BGP, enters the configuration.

To reduce the risk of erroneous route advertising, the ISP should always filter any BGP information received from the customer network. The ISP should reject routes to networks that are not expected to be in the customer AS. Routes that contain an AS path with unexpected AS numbers should also be rejected.

Example


Configuring BGP on Multihomed Customer Routers (Cont.)
Cisco.com
ip route 11.2.3.0 255.255.255.0 null 0
router bgp 65001
neighbor 11.1.1.1 remote-as 65001
neighbor 10.0.0.1 remote-as 387
network 11.2.3.0 mask 255.255.255.0
router ospf 1
default-information originate
IGP session
BGP
BGP
11.2.3.0/24
AS 65001
Customer Network
AS 387
Service Provider Network
• Customer address space is advertised on every customer edge router
• Customer edge routers run IBGP between themselves and advertise default route to the rest of the customer network
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-11

In this figure, the customer has been assigned the private AS number 65001. Both customer edge routers are configured to run BGP and should advertise all of the customer networks with the **network** command. If only one router advertises the network, a single point of failure has been introduced. The two customer edge routers must also run Internal Border Gateway Protocol (IBGP) between them to make common decisions regarding BGP routing information.

Each customer edge router has an External Border Gateway Protocol (EBGP) session with the ISP router on the other side of the link. Over that EBGP session, the ISP announces only a default route to the customer AS. When EBGP receives the default route, it installs it in the forwarding table and redistributes it into the Interior Gateway Protocol (IGP) (in this case Open Shortest Path First [OSPF]) of the customer.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Why can you not use static routing in all cases of redundant links between a customer and a single ISP?

A) You cannot use static routing when multihoming to a single provider.

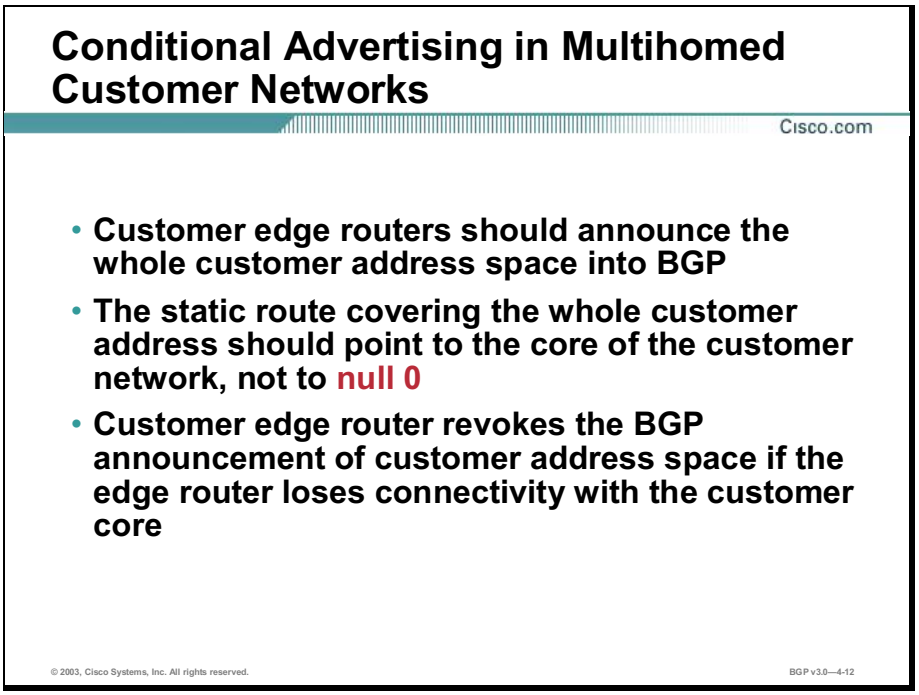
B) You cannot use static routing when load balancing is a design requirement.

C) You cannot use static routing when the ISP conditionally advertises customer routes.

D) You cannot use static routing in scenarios where the failures cannot be detected by Layer 2 protocols.

Conditional Advertising in Multihomed Customer Networks

This topic describes how to configure conditional advertising of a customer address space when you are using BGP.


Conditional Advertising in Multihomed Customer Networks
Cisco.com
• Customer edge routers should announce the whole customer address space into BGP
• The static route covering the whole customer address should point to the core of the customer network, not to null 0
• Customer edge router revokes the BGP announcement of customer address space if the edge router loses connectivity with the customer core
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-12

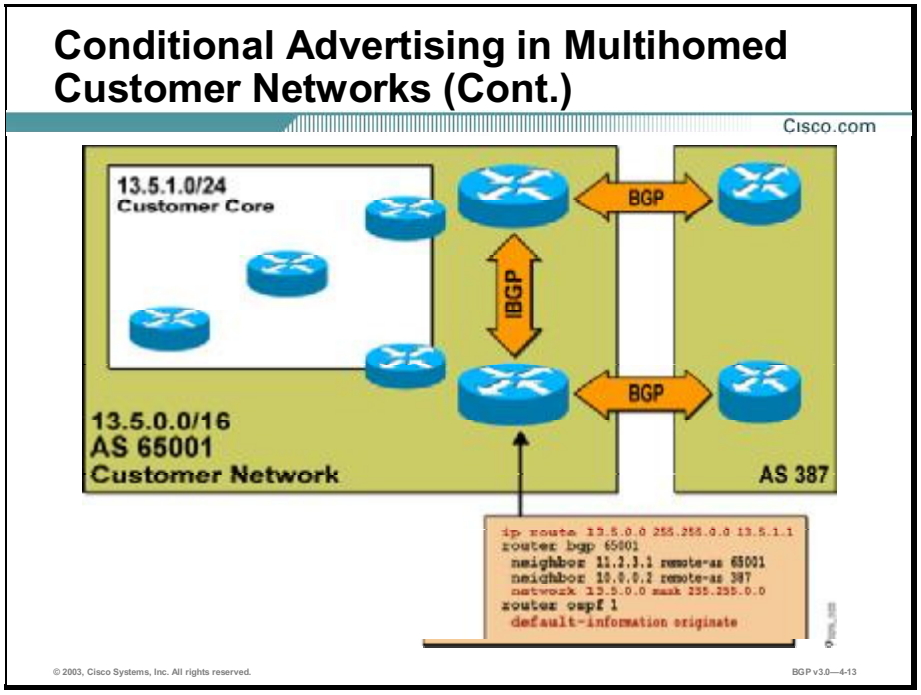
As a rule of thumb, the customer should announce addresses as large as possible (the larger the address space that can be aggregated, the better). The BGP advertisement is configured on the customer edge routers using the **network** command. Route advertisement is conditioned by the appearance of a corresponding network or subnet in the forwarding table of the edge router. If the network or subnet is manually entered into the forwarding table by configuring a static route to “null 0,” the condition is always true because the static route is always there, and the BGP advertisement is always performed.

If the customer edge router loses connectivity to the rest of the customer network but is still connected to the ISP network, BGP advertisement must cease. In this case, BGP advertisement can be stopped if BGP advertisements are bound to the reachability status of a specific subnet in the core of the customer network, according to the customer IGP.

The problem with using a static route to null0 is that it conditions the network statement in the BGP configuration so that BGP always advertises the route. If the customer edge router loses connectivity with the rest of the customer network, the router continues to advertise the entire customer address space. The ISP network receives a valid route from the customer edge router. Traffic is sent to this router, but because it has lost connectivity with the rest of the network, the traffic is dropped (routed to the null0 interface using the static route).

 Copyright © 2003, Cisco Systems, Inc.

Example



In this example, the customer network uses the address space 13.5.0.0/16. The address space is further subnetted at the customer site. One of the subnets (subnet 13.5.1.0/24) is identified as being a central part of the customer core network.

The customer edge routers participate in the IGP routing of the customer. This participation means that these routers have information about which of the subnets within the address space 13.5.0.0/16 are currently reachable. If these subnets are available, there is an explicit route to each of them. If any of the subnets go down, or the path toward them goes down, the route to that subnet is removed from the forwarding table.

The BGP advertisement in each of the customer edge routers is configured to advertise the full address space used by the customer. When this route is advertised by the customer edge routers, the ISP network, and thus the rest of the Internet, see the complete address space of the customer as one single route, 13.5.0.0/16.

Advertisement of the customer address space by BGP is conditioned by the appearance of the static route, **ip route 13.5.0.0 255.255.0.0 13.5.1.1**. If the static route is valid, then the BGP route 13.5.0.0/16 is advertised. The static route is a recursive route, which means that the router takes another look in the routing table for the address 13.5.1.1 before determining what to do with the static route. The idea is that 13.5.1.1 is reachable via the IGP. The subnet 13.5.1.0/24 is announced by the IGP. If this subnet is reachable by the edge router, then the static route to 13.5.0.0/16 is valid. If there is no route to 13.5.1.1, then the static route is invalid.

Note	The condition, whether or not to advertise the entire customer address space 13.5.0.0/16, is controlled by the IGP reachability of a single subnet, 13.5.1.0/24.

The IGP configuration also includes origination of the default route by both edge routers.

Practice

Q1) Why should you not configure conditional route advertising using a route to null 0?

A) Because the route to null 0 will cause all traffic to de discarded.

B) A route to null 0 will condition BGP to advertise routes only when the interface with a matching assigned network number is in the up state.

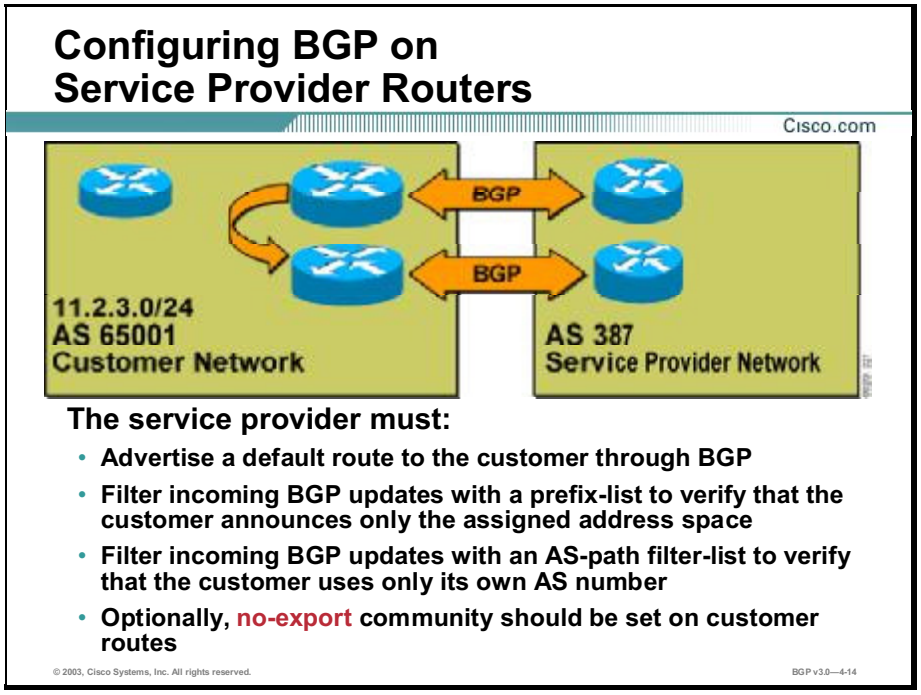
C) A route to null 0 will condition BGP such that the network will always be advertised regardless of its state.

D) Using null 0 as a conditional advertisement will advertise the network only if a matching route exists in the IP routing table.

 Copyright © 2003, Cisco Systems, Inc.

Configuring BGP on Service Provider Routers

This topic describes how to configure BGP on a service provider network to establish routing between a multihomed customer and a single service provider.

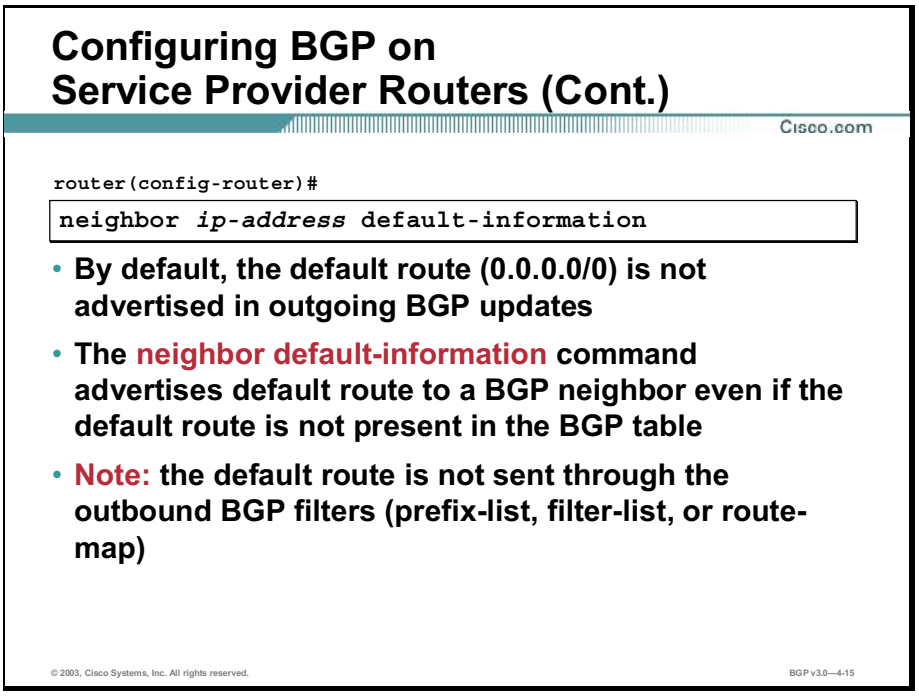

Configuring BGP on Service Provider Routers
Cisco.com
BGP
BGP
11.2.3.0/24
AS 65001
Customer Network
AS 387
Service Provider Network
The service provider must:
• Advertise a default route to the customer through BGP
• Filter incoming BGP updates with a prefix-list to verify that the customer announces only the assigned address space
• Filter incoming BGP updates with an AS-path filter-list to verify that the customer uses only its own AS number
• Optionally, no-export community should be set on customer routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-14

In the ISP network, the two edge routers must have BGP sessions configured for the customer. There is no point in feeding the full Internet routing table to the customer, because the table contains the same set of routes for both links and the customer always uses the ISP for all traffic toward the Internet. A default route injected in the customer network would accomplish the same task.

The customer is responsible for its own advertisements. Assuming that the customers are much less experienced in BGP configuration than the ISP, they are more likely to make errors. Therefore, the ISP must protect itself and the rest of the Internet from those errors.

The service provider should use a prefix-list that allows only customer-assigned routes and denies any other route to ensure that private address space or any other illegal networks erroneously announced by the customer never reach the ISP BGP table. Filtering based on the AS path also provides some protection from customer configuration errors. Only routes originated within the customer AS are allowed. A filter-list performs this check.

If the customer address space is PA address space, and it represents only a small part of a larger block announced by the ISP, the explicit BGP routes received from the customer need not be advertised to the rest of the Internet. The ISP can announce the big block, attracting any traffic toward any subnet within the block. After the traffic enters the ISP network, the more explicit routes to the customer network are available and used. In this case, the provider edge router can tag the BGP routes received from the customer with the **no-export** well-known community, restricting them from being sent by the ISP to any other AS.

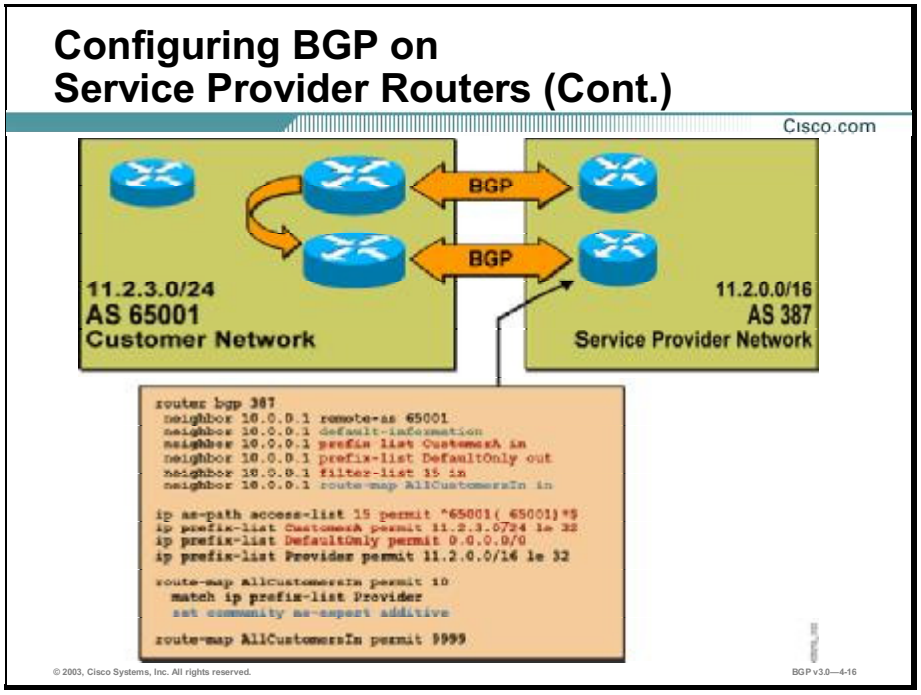

Configuring BGP on
Service Provider Routers (Cont.)
Cisco.com
router(config-router)#
neighbor ip-address default-information
• By default, the default route (0.0.0.0/0) is not advertised in outgoing BGP updates
• The neighbor default-information command advertises default route to a BGP neighbor even if the default route is not present in the BGP table
• Note: the default route is not sent through the outbound BGP filters (prefix-list, filter-list, or route-map)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-15

The default route, 0.0.0.0/0, is not advertised in outgoing BGP updates unless it is explicitly configured. The **neighbor default-information** router configuration command is used to initiate the advertisement of the default route to a neighbor.

No checking is done by BGP before the default route is advertised. The default route does not need to be present in the BGP table before it is advertised using this command. The default route is also sent without being filtered by any outgoing prefix-lists, filter-lists, or route-maps.

 Copyright © 2003, Cisco Systems, Inc.

Example



This example shows the configuration of an ISP edge router.

The customer is assigned the private AS number 65001. The BGP session is opened with the customer IP address, 10.0.0.1.

The ISP sends the default route only to the customer. This route is configured first using first the **default-information** command and then the prefix-list *Default-only*.

Received routes from the customer must first pass the prefix-list *CustomerA*. There is one dedicated prefix-list for each individual customer, permitting only those routes that the customer is allowed to announce. If the routes are allowed by the prefix-list, they must also pass the filter-list *15 in*. In this case, the filter allows the private AS of the customer in any number of repetitions, as long as it is the only AS number in the path. This filter-list allows for AS-path prepending configurations on the customer side. If the received route is allowed by both the prefix-list and the filter-list, then the route-map *AllCustomersIn* is applied.

The route-map is a general route-map that is used for all customers. This map checks every route received, via the prefix-list *Provider*, and if the route is within the big block of the PA address space, which the ISP announces to the rest of the Internet, the customer route is marked with the **no-export** community. This mark means that the route is used within the ISP AS only, and is not sent to the rest of the Internet.

Routes that are received from the customer, and are allowed by the prefix-list and filter-list, but do not fall within the PA address space, are allowed by the route-map and are not changed in any way. The ISP propagates these routes to the rest of the Internet.

Practice

Q1) Functionally, what three requirements must you take into consideration when configuring in the service provider network to support a multihomed customer? (Choose three.)

A) The provider should filter incoming BGP updates with a prefix-list to verify customer-announced assigned space.

B) The provider should filter incoming BGP updates with an AS-path filter-list to verify that the customer uses its own AS number only.

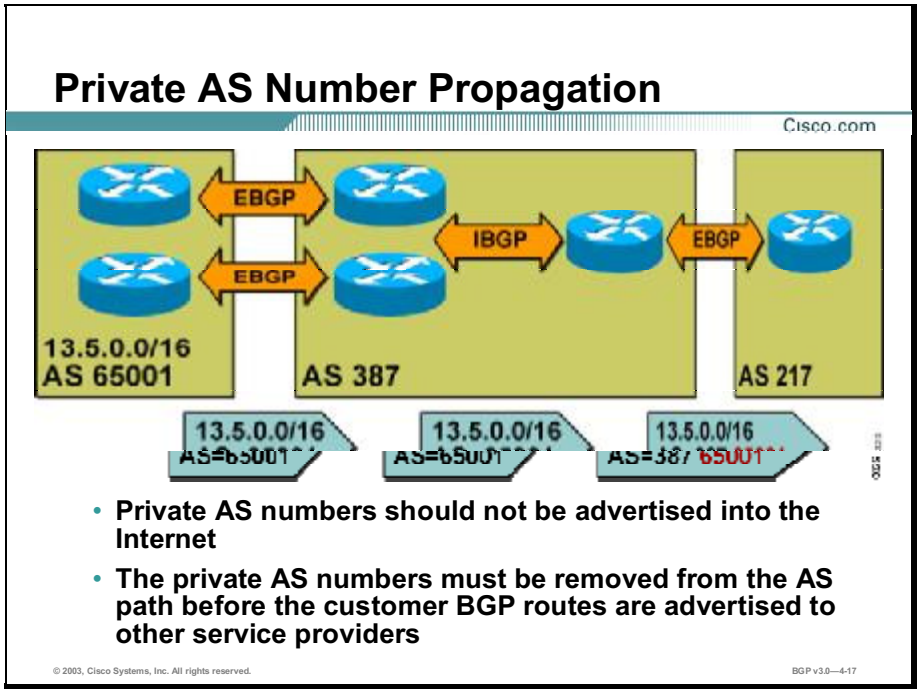
C) The provider should advertise a default route to the customer through BGP.

D) The provider should announce nonsummarized prefixes for the customer networks to the Internet.

 Copyright © 2003, Cisco Systems, Inc.

Removing Private AS Numbers

This topic describes how to disable the propagation of private AS numbers to EBGP peers in a service provider network where a multihomed customer is advertising private numbers in the AS path.

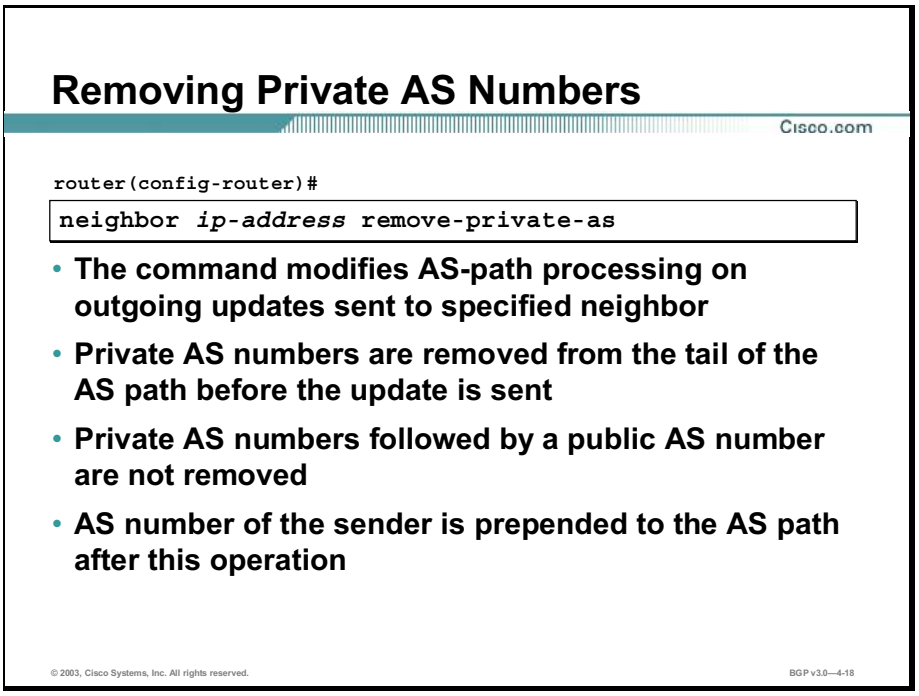


Routes received by the ISP from the customer are propagated to the rest of the Internet only if they are part of the provider-independent (PI) address space.

When the ISP receives BGP routes from the customer, the AS-path attribute of the received routes contains only the AS number of the customer. If the customer uses AS-path prepending, there may be several repetitions of the customer AS number in the AS-path. If customer routes are propagated by the service provider to the Internet, the AS number of the customer will be present in the AS path unless explicitly removed.

Note	If the customer has been assigned a private AS number, this AS number must never be advertised by any router to the rest of the Internet.

Removal of a private AS number from the AS path is accomplished by using **remove-private-as** on the ISP EBGP sessions with the rest of the Internet. In the figure, removal of the private AS number takes place on the EBGP session between AS 387 and AS 217.


Removing Private AS Numbers
Cisco.com
router(config-router)#
neighbor ip-address remove-private-as
• The command modifies AS-path processing on outgoing updates sent to specified neighbor
• Private AS numbers are removed from the tail of the AS path before the update is sent
• Private AS numbers followed by a public AS number are not removed
• AS number of the sender is prepended to the AS path after this operation
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-18

neighbor remove-private-as

To remove private AS numbers from the AS path (a list of AS numbers that a route passes through to reach a BGP peer) in outbound routing updates, use the **neighbor remove-private-as** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **remove-private-as**

To disable this function, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **remove-private-as**

Syntax Description

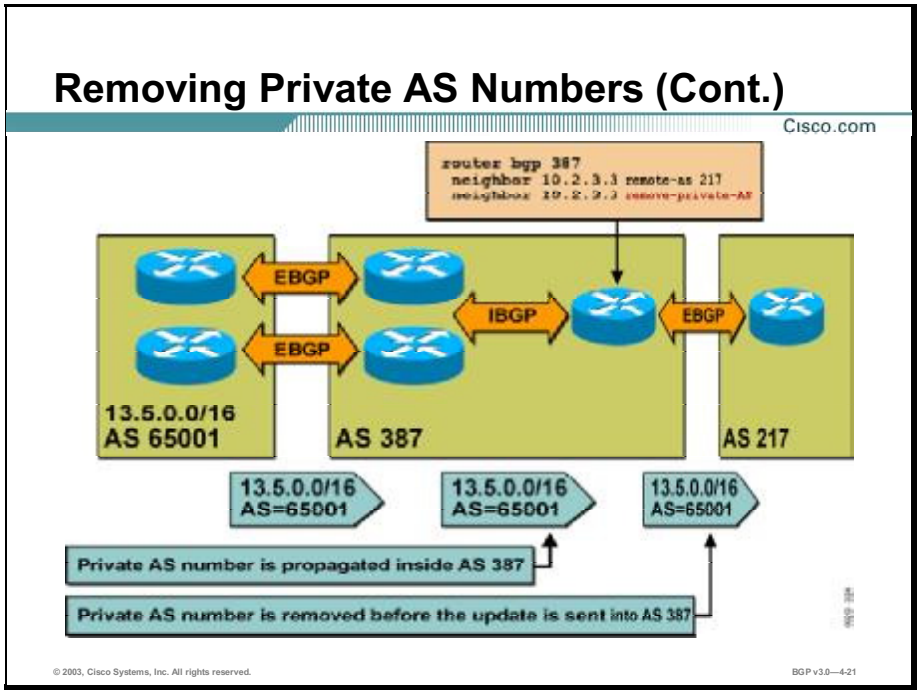
ip-address	IP address of the BGP-speaking neighbor
peer-group-name	Name of a BGP peer group

Use this command on the service provider egress routers. Before any of the customer routes of the ISP are advertised by the service provider to the rest of the Internet, the AS numbers in the range 64512-65535 must be removed. The command removes those AS numbers if they are in the tail end of the AS path. Private AS numbers followed by public AS numbers are not removed.

The AS number of the ISP is automatically prepended to the AS-path attribute after the **remove-private-as** operation has completed. This situation means that the AS number of the ISP has not already been prepended to the AS-path attribute when the tail of the AS path is checked for private AS numbers.

 Copyright © 2003, Cisco Systems, Inc.

Example



In this example, the service provider AS (387) receives routes from the customer. The customer is assigned the private AS number 65001 by the ISP. Therefore, routes received by the provider have an AS path containing only AS 65001. This information should be kept and used within the ISP network and should never be propagated to the rest of the Internet (AS 217 in this example).

The edge router in AS 387 has been configured to remove private AS numbers on EBGP routes toward AS 217. If private AS numbers appear in the tail end of the AS path (before AS 387 is added), they are removed.

This configuration must be applied to all egress router in AS 387 that contain EBGP neighbors leading to other ISPs. No private AS number may be present in an AS path of a route propagated to a network using a public AS number.

Practice

Q1) AS 64525 is connected to AS 229, which in turn is connected to AS 1126. How will configuring private AS removal on AS 229 affect routers in AS 1126?

A) There will be no effect on the routers in AS 1126.

B) The routers in AS 1126 will see all Internet routes as originating in AS 64525.

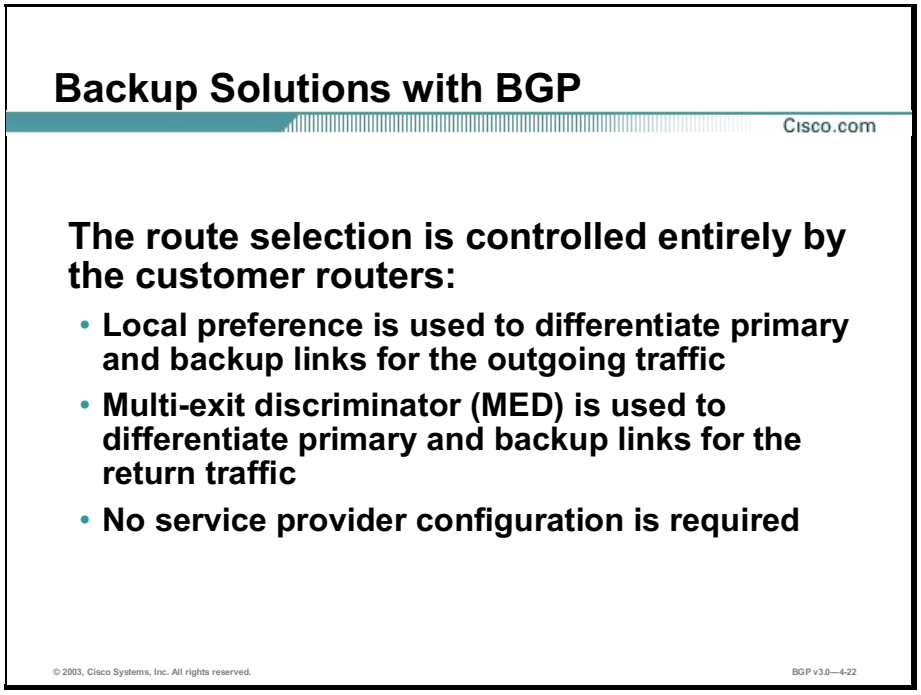
C) The routers will see networks originating in AS 64525 as originating in AS 229.

D) AS 1126 will receive all of its routes with an AS-path length of two.

 Copyright © 2003, Cisco Systems, Inc.

Backup Solutions with BGP

This topic describes how to configure a typical backup setup between a multihomed customer and a service provider in a BGP environment.


Backup Solutions with BGP
Cisco.com
The route selection is controlled entirely by the customer routers:
• Local preference is used to differentiate primary and backup links for the outgoing traffic
• Multi-exit discriminator (MED) is used to differentiate primary and backup links for the return traffic
• No service provider configuration is required
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-22

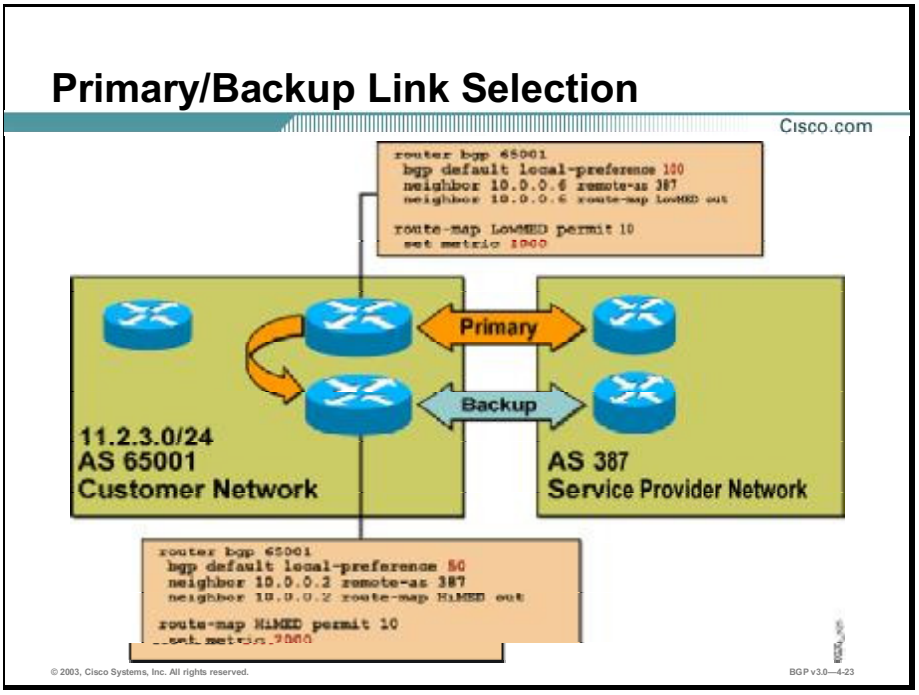
When a customer uses BGP on multiple links between its network and the ISP network, the customer is solely responsible for controlling how it uses the links. The customer can choose to use its links in a primary/backup scenario or in a load-sharing scenario.

If one link is primary, then the other should be used for backup only. The customer can use the local preference configuration to direct all outgoing traffic over the primary link.

Incoming traffic to the customer is controlled using either AS-path prepending or multi-exit discriminator (MED). Because the customer has multiple connections to the same AS, MED is the ideal attribute to use. When the customer announces its routes to the ISP, a bad (high) MED value on the backup link and a good (low) value on the primary link are set.

MED (and AS-path length) is checked by the receiving EBGP peer only if the weight and local preference attributes have not been configured. In this case, the ISP should not use any of these configuration options. The ISP should rely solely on the attributes received from the customer.

Example



In the figure, the customer is connected to the ISP over two permanent connections. The customer uses the upper connection as the primary connection and the lower connection as the backup.

The BGP configuration on the ISP side is transparent. This transparency means that no particular preference is configured to use the upper or lower connection. The ISP relies on the attribute values received from the customer.

The primary edge router on the customer side is configured to set local preference to the value 100 on all EBGP routes received. The backup edge router sets the local preference attribute to a value of 50. This configuration means that the outgoing traffic toward any destination announced by the ISP is primarily sent over the upper link.

Incoming traffic to the customer is directed to the primary link using the MED. In the primary edge router of the customer, all routes that are sent to the ISP have their MED attribute set to the value 100 by the route-map *LowMED out*. In the backup edge router of the customer, all routes that are sent to the ISP have their MED attribute set to the value 2000 by the route-map *HiMED out*. Because the ISP receives the routes with all other attributes set to the same values, the MED values direct traffic, destined to the customer, to the primary link.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Which attribute can you use to select the primary/backup link for outgoing traffic?

A) weight

B) local preference

C) AS-path

D) MED

Q2) Which attribute can you use to select the primary/backup link for incoming traffic?

A) weight

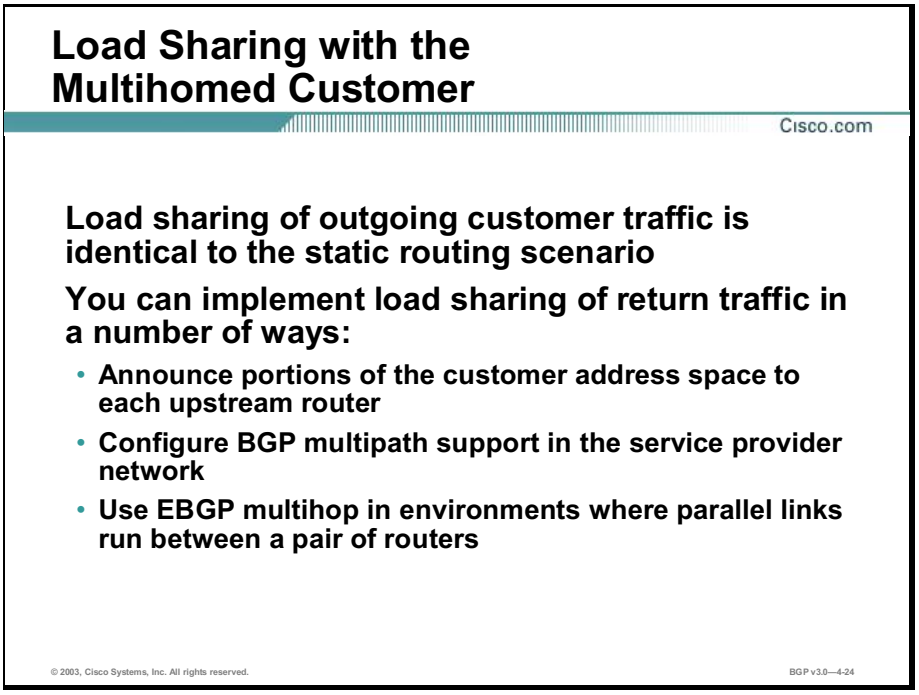
B) local preference

C) AS-path

D) MED

Load Sharing with the Multihomed Customer

This topic describes how you can implement load sharing between a multihomed customer and a service provider in a BGP environment.


Load Sharing with the Multihomed Customer
Cisco.com
Load sharing of outgoing customer traffic is identical to the static routing scenario
You can implement load sharing of return traffic in a number of ways:
• Announce portions of the customer address space to each upstream router
• Configure BGP multipath support in the service provider network
• Use EBGP multihop in environments where parallel links run between a pair of routers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-24

Load sharing of outgoing traffic from the customer network is identical to the static routing scenario. The customer IGP is configured to send information about a gateway of last resort. There is no difference whether the edge router gets its default by static routing or by incoming EBGP updates.

Load sharing of the return traffic coming back to the customer network from the ISP can be implemented in a number of ways:

- The customer could divide its address space into several announcements. The customer edge router can send each announcement over one of its EBGP sessions with the ISP. For backup purposes, the customer should advertise the entire address space over all of its EBGP sessions. The ISP now uses the most explicit route rule, and as long as both links are up, traffic with destinations within one part of the customer address space is routed over one of the links and traffic to the other part is routed over the other link.

- If the customer announces equivalent routes over both links, the ISP routers use the closest connection with respect to the IGP of the ISP. If an ISP router has an equivalent distance to both connection points, the use of the maximum-paths (BGP multipath) option causes load sharing.

 Copyright © 2003, Cisco Systems, Inc.

- If the multiple links between the customer and the ISP network terminate in one single router on the customer side and one single router on the ISP side, the two routers must establish their EBGP session from loopback interface to loopback interface. Static or dynamic routing is required for one router to get information on how to reach the loopback interface of the other router. The use of the **ebgp-multihop** option is also required because the address of the neighbor is not directly connected.

Practice

Q1) What three options can you use to enable load sharing on parallel links connected to one router? (Choose three.)

A) Split the customer address space into two parts and advertise a portion on each link.

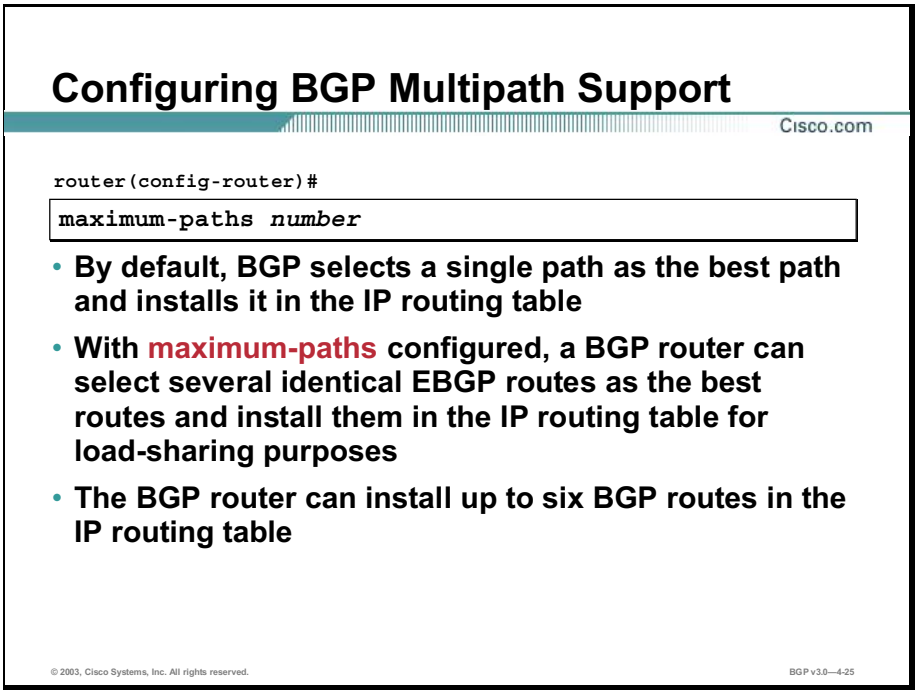
B) Use AS-path prepending on the outgoing routes of the backup path.

C) Use the **ebgp-multihop** option between loopback interfaces of the multihomed routers.

D) Enable BGP multipath support on the multihomed routers.

Load Sharing with BGP Multipath

This topic presents the Cisco IOS command required to configure load sharing between a multihomed customer and a service provider through the use of BGP multipath.


Configuring BGP Multipath Support
Cisco.com
router(config-router)#
maximum-paths number
• By default, BGP selects a single path as the best path and installs it in the IP routing table
• With maximum-paths configured, a BGP router can select several identical EBGP routes as the best routes and install them in the IP routing table for load-sharing purposes
• The BGP router can install up to six BGP routes in the IP routing table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-25

By default, BGP route selection rules select one, and only one, route as the best. If there are two identical routes, the tiebreaker is the most stable route or the router-ID of the peer router advertising the route. However, by using the **maximum-paths** router configuration command, the BGP route selection process will select more than one route as best if they are exactly identical. The routes are all installed in the forwarding table, and load sharing will take place.

maximum-paths

To control the maximum number of parallel routes that an IP routing protocol can support, use the **maximum-paths** command in address family or router configuration mode.

maximum-paths *number*

To restore the default value, use the **no** form of this command

no maximum-paths

 Copyright © 2003, Cisco Systems, Inc.

Syntax Description

number	Maximum number of parallel routes that an IP routing protocol installs in a routing table, in the range of 1 to 6

Note Load sharing between alternative BGP routes is achieved only if the EBGP routes are identical according to all BGP route selection rules and **maximum-paths** is configured with a value larger than 1.

A BGP router can install up to six BGP routes in the IP forwarding table. The actual type (per-session or per-packet) of load sharing done between the routes depends on the switching mode used.

Practice

Q1) By default, BGP can perform load balancing over how many parallel links?

A) one

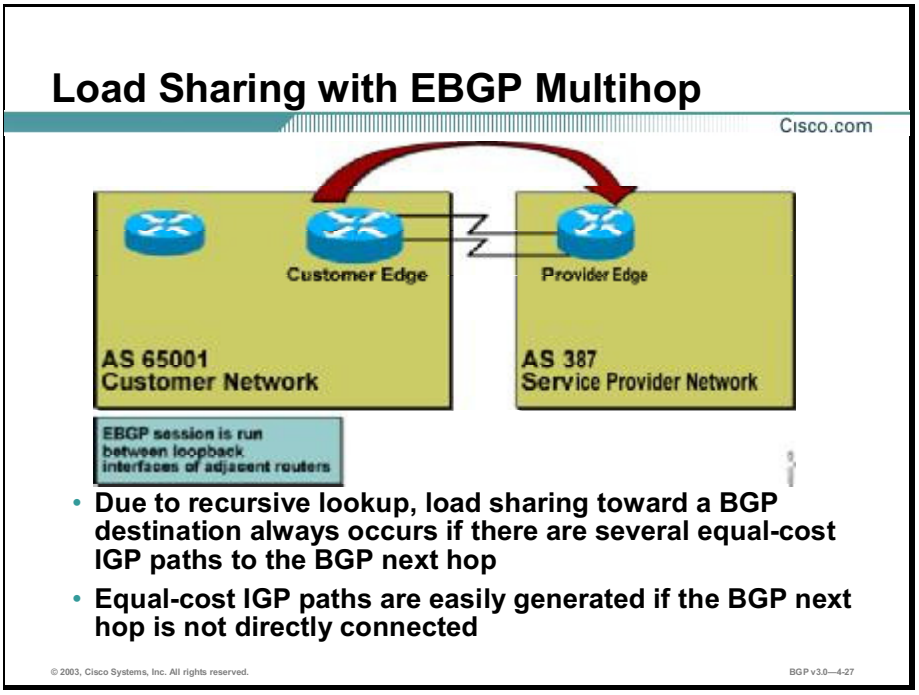
B) four

C) six

D) eight

Load Sharing with EBGP Multihop

This topic describes how to configure load sharing between a multihomed customer and a service provider through the use of EBGP multihop.


Load Sharing with EBGP Multihop
Cisco.com
Customer Edge
Provider Edge
AS 65001
Customer Network
AS 387
Service Provider Network
EBGP session is run between loopback interfaces of adjacent routers
• Due to recursive lookup, load sharing toward a BGP destination always occurs if there are several equal-cost IGP paths to the BGP next hop
• Equal-cost IGP paths are easily generated if the BGP next hop is not directly connected
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-27

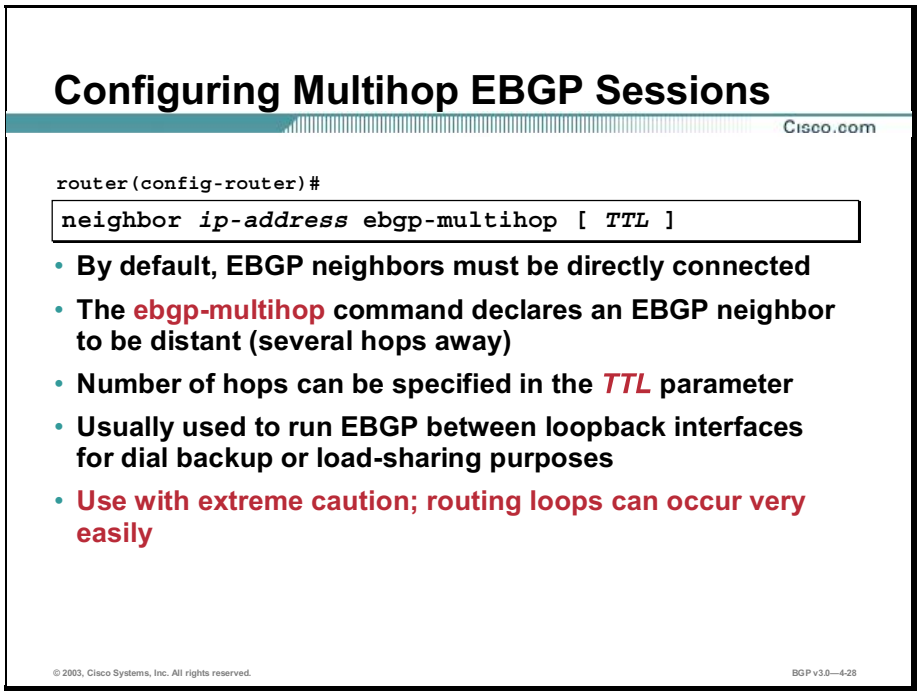
When two adjacent routers have multiple links between them, you can configure the EBGP session from loopback interface to loopback interface. In this case, you must use **ebgp-multihop** to make the BGP session go into the active state. There must be static or dynamic routing in use to provide both routers with information on how to reach the loopback interfaces of each other. Otherwise, their EBGP session does not complete establishment.

Routing to the loopback interface of the neighboring router is required to establish the EBGP session and is also used in the recursive lookup when the routes are installed by the router in its forwarding table. The two routes to the loopback interface of the neighboring router should be equivalent for load sharing to occur.

After configuration, one single EBGP session is established between the two routers. This session is used to exchange the routing information. There is only one BGP route to each destination, and it has a next hop referring to the loopback interface of the other router.

Before installing a route to a specific destination in its forwarding table, a router will perform a recursive lookup to resolve the next hop. In this case, the recursive lookup will result in finding two alternative routes. The router will install the BGP route to the final destination twice in the forwarding table (Forwarding Information Base [FIB]). The first time, the route is installed with one of the resolved next-hop addresses, and the second time with the other resolved next-hop address. Since multiple equal-cost paths exist, the router can load share over the two paths depending on the switching mode.

 Copyright © 2003, Cisco Systems, Inc.

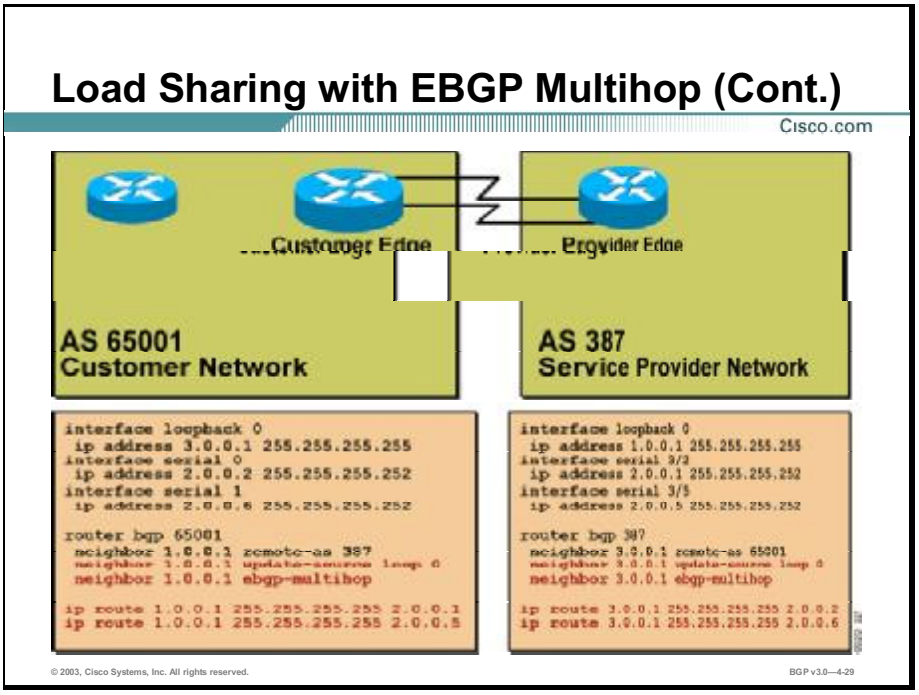

Configuring Multihop EBGP Sessions
Cisco.com
router(config-router)#
neighbor ip-address ebgp-multihop [TTL]
• By default, EBGP neighbors must be directly connected
• The ebgp-multihop command declares an EBGP neighbor to be distant (several hops away)
• Number of hops can be specified in the TTL parameter
• Usually used to run EBGP between loopback interfaces for dial backup or load-sharing purposes
• Use with extreme caution; routing loops can occur very easily
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-28

By default, EBGP neighbors must be directly connected. Cisco IOS software verifies that an EBGP neighbor is reachable as directly connected over one of the router interfaces before the session goes into the active state. For an EBGP session, IP packets carrying the TCP segments with BGP information are also sent using a Time to Live (TTL) value set to the value 1. This value means that they cannot be routed.

The **ebgp-multihop** neighbor configuration command changes this behavior. Although the neighbor is several hops away, the session goes into the active state, and packets start to be exchanged. The TTL value of the IP packets is set to a value larger than 1.If no value is specified on the command line, 255 is used.

Use the **ebgp-multihop** command when establishing EBGP sessions between loopback interfaces for load-sharing purposes. You must take great care when using **ebgp-multihop**, because proper packet forwarding relies on all the intermediate routers along the path to the EBGP peer to make the correct forwarding decision. If the intermediate routers have a correct path to the EBGP peer, but a wrong path to the final destination, the packet may get into a routing loop.

Example



In the figure, the customer network and the ISP network are connected using two parallel links between a single router on the customer side and a single router on the ISP side.

In this case, only one EBGP session is configured between the customer and provider routers. The session should be established from the loopback interface in one router to the loopback interface in the other.

Each of the two edge routers has two static host routes pointing to the loopback interface on the other router. The EBGP session is established from loopback to loopback using **ebgp-multihop**.

The customer receives an EBGP route from the ISP with the next hop set to 1.0.0.1. The customer edge router performs a recursive lookup and finds that it can reach 1.0.0.1 via 2.0.0.1 and via 2.0.0.5. These two routes are equivalent. Therefore, the route to the final destination is installed in the forwarding table of the customer router using both paths.

Depending on the switching mode in use, load sharing is done per packet, per destination, or per source/destination pair.

In this example, link-level procedures ensure that if one of the links goes down, the corresponding static link goes down. All BGP routes in the forwarding table that rely on the static route to the link that went down are invalidated. However, the BGP routes in the forwarding table relying on the remaining link are still valid and used.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What three criteria must be met before two routers with parallel links between them can perform load balancing using **ebgp-multihop**? (Choose three.)

A) The **neighbor ebgp-multihop** command must be configured on each router.

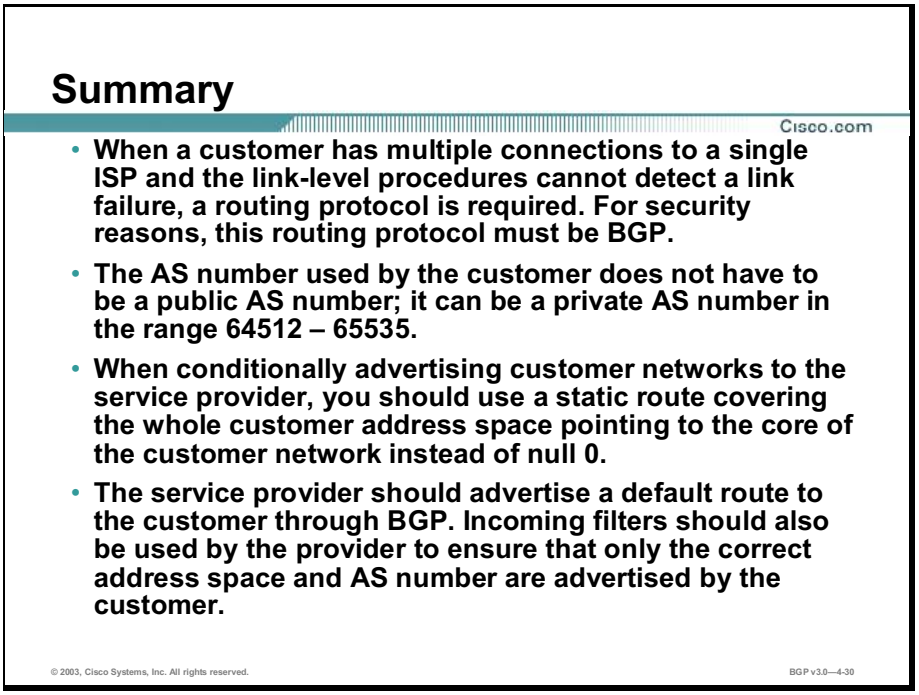
B) BGP maximum paths must be set to the number of links between the routers.

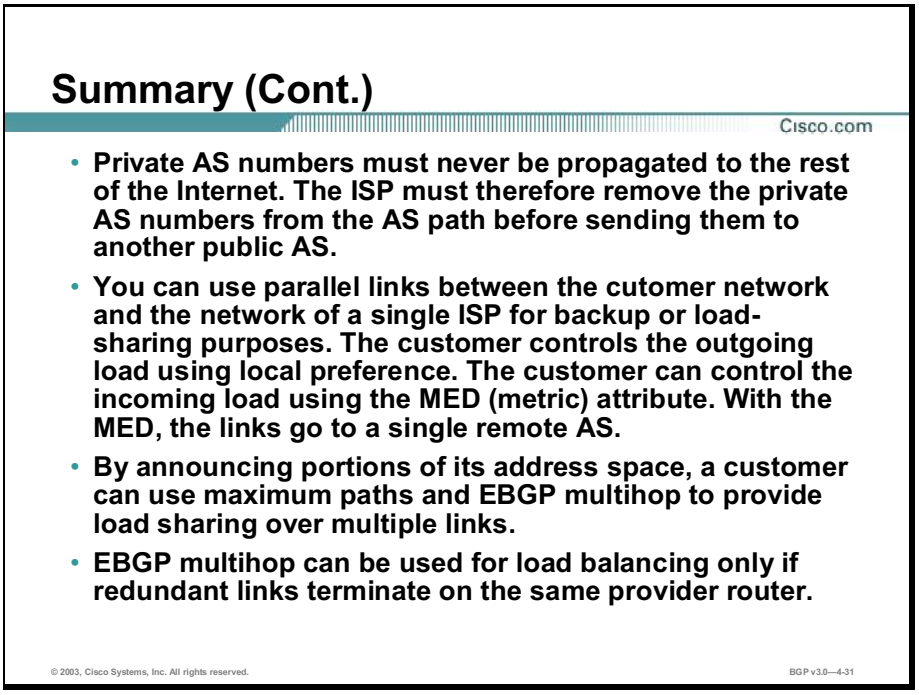
C) The routers must have a static route or an IGP containing reachability information for the configured loopback addresses.

D) The **neighbor update-source** option must be configured on both routers.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• When a customer has multiple connections to a single ISP and the link-level procedures cannot detect a link failure, a routing protocol is required. For security reasons, this routing protocol must be BGP.
• The AS number used by the customer does not have to be a public AS number; it can be a private AS number in the range 64512 – 65535.
• When conditionally advertising customer networks to the service provider, you should use a static route covering the whole customer address space pointing to the core of the customer network instead of null 0.
• The service provider should advertise a default route to the customer through BGP. Incoming filters should also be used by the provider to ensure that only the correct address space and AS number are advertised by the customer.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-30


Summary (Cont.)
Cisco.com
• Private AS numbers must never be propagated to the rest of the Internet. The ISP must therefore remove the private AS numbers from the AS path before sending them to another public AS.
• You can use parallel links between the cutomer network and the network of a single ISP for backup or load-sharing purposes. The customer controls the outgoing load using local preference. The customer can control the incoming load using the MED (metric) attribute. With the MED, the links go to a single remote AS.
• By announcing portions of its address space, a customer can use maximum paths and EBGP multihop to provide load sharing over multiple links.
• EBGP multihop can be used for load balancing only if redundant links terminate on the same provider router.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—4-31

 Copyright © 2003, Cisco Systems, Inc.

Next Steps

After completing this lesson, go to:

- Connecting a Multihomed Customer to Multiple Service Providers lesson

References

For additional information, refer to these resources:

- For more information on removing private AS numbers, refer to “Removing Private Autonomous System Numbers in BGP” at the following URL: http://www.cisco.com/warp/public/459/32.html

- For more information on using MED for path selection, refer to “How BGP Routers Use the Multi-Exit Discriminator for Best Path Selection” at the following URL: http://www.cisco.com/warp/public/459/37.html

- For more information on load sharing, refer to “Sample Configurations for Load Sharing with BGP in Single and Multihomed Environments” at the following URL: http://www.cisco.com/warp/public/459/40.html

Copyright © 2003, Cisco Systems, Inc.

Quiz: Connecting a Multihomed Customer to a Single Service Provider

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Configure BGP on a customer network to establish routing between a multihomed customer and a single service provider
- Configure conditional advertising of a customer address space when BGP is used to establish routing between a multihomed customer and a single service provider
- Configure BGP on a service provider network to establish routing between a multihomed customer and a single service provider
- Disable the propagation of private AS numbers to EBGP peers in a service provider network where a multihomed customer is advertising private numbers in the AS path
- Configure a typical backup setup between a multihomed customer and a single service provider in a BGP environment
- Describe how you can implement load sharing between a multihomed customer and a single service provider in a BGP environment
- Identify the Cisco IOS command required to configure load sharing between a multihomed customer and a single service provider using BGP multipath
- Configure load sharing between a multihomed customer and a single service provider using EBGP multihop

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

 Copyright © 2003, Cisco Systems, Inc.

Q1) What are three responsibilities of the customer when the customer is multihomed to a single service provider? (Choose three.)

A) Customer edge routers must run IBGP between them.

B) The customer must advertise a default route.

C) The customer must conditionally advertise its assigned address space into BGP.

D) The customer edge routers must run EBGP with the provider.

Q2) Given the following router command output, what method has been used to influence return traffic in a primary/backup link implementation for this multihomed customer?

```
Provider# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.10.20.0/24    192.168.63.3          1000             0 100 100
i
*>                  192.168.64.4          2000             0 400 100
i
*> 30.30.30.0/24    192.168.63.3             0             0 100 I
*> 40.40.40.0/24    192.168.64.4             0             0 400 I
```

A) MED

B) local preference

C) weight

D) AS-path prepending

Q3) What are three responsibilities of the provider router when supporting a multihomed customer? (Choose three.)

A) The provider must advertise a default route to the customer through BGP.

B) The provider must filter customer routes to verify that proper addressing is used.

C) The provider must remove the private AS number, if in use by the customer.

D) The provider must configure new AS-path filters to allow AS-path prepending; otherwise, a primary/backup link cannot be established.

Q4) What will occur if private AS numbers are advertised to the Internet?

A) The Internet will not be able to route packets.

B) Internet routers could drop routes based on BGP loop prevention mechanisms.

C) Customer load balancing will not function.

D) Customer configurations for the primary/backup link using AS-path prepending will not function.

Q5) What BGP configuration is required to properly implement a backup solution for a multihomed customer connected to a single provider? (Choose two.)

A) The customer should set local preference to influence outgoing route selection.

B) The customer should set the weight attribute to influence outgoing path selection.

C) The customer should set MED on each route to influence return path selection.

D) The customer should configure AS-path prepending to ensure proper outgoing path selection.

Q6) A customer router has been configured with maximum paths set to a value of 4. Given the following router command output, over how many links will the router need to perform load balancing?

```
router# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.10.20.0/24    192.168.63.3                           0 300 100
100 i
*>                  192.168.64.4                           0 400 100
i
*                   192.168.65.5                           0 500 100
i
*> 30.30.30.0/24    192.168.63.3             0             0 300 i
*> 40.40.40.0/24    192.168.64.4             0             0 400 i
```

 Copyright © 2003, Cisco Systems, Inc.

A) The router will use only the path marked as “best” by BGP.

B) The router will perform load balancing over two paths to reach network 10.10.20.0/24.

C) The router will perform load balancing over three paths to reach network 10.10.20.0/24.

D) There is not enough information to determine the correct answer.

Q7) What three methods can you use to provide load sharing over network links between a multihomed customer and a single provider? (Choose three.)

A) advertising of split addressing space to the provider

B) configuring ebgp-multihop between the customer and the provider

C) use of the BGP **maximum-paths** command to perform load balancing over parallel links

D) configuring multiple static routes pointing to the provider

Q8) Why is it not required to configure maximum paths under the BGP routing process when you are performing load balancing with ebgp-multihop?

A) By default, BGP will perform load balancing over up to four paths, configurable up to six.

B) The static route or IGP process is responsible for load balancing in this configuration.

C) Configuring multihop enables maximum paths equal to the TTL setting of the **neighbor ebgp-multihop** command.

D) Configuring ebgp-multipath is a required component of ebgp-multihop load balancing.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Connecting a Multihomed Customer to Multiple Service Providers

Overview

Customers connect to the Internet using service providers to enable different applications such as intranet connectivity with Virtual Private Networks (VPNs), extranet connectivity with suppliers, and other Internet applications. Different customers have different connectivity requirements depending upon their business model, redundancy requirements, and even network budget.

This lesson discusses using multiple connections between a customer and multiple service providers for backup and load-sharing purposes. Included in this lesson is a discussion of the Border Gateway Protocol (BGP) characteristics used to configure customer and provider networks to accommodate the multiple connections between them. Also discussed in this lesson are topics specific to networks with multiple connections between a customer and multiple providers such as address selection, private autonomous system (AS) number translation, and configuration of the network to support either backup links or load sharing.

Importance

When a customer requires the maximum redundancy in its network design, it should implement a multihomed strategy that uses multiple service providers. This configuration requires specific considerations to be implemented properly. Addressing and AS number selection are important considerations that affect the implementation of the network. It is also important to understand how to configure routing protocols so that customer backup or load-sharing requirements are met.

Objectives

Upon completing this lesson, you will be able to:

- Describe BGP configuration characteristics used to establish routing between a multihomed customer and multiple service providers
- Describe addressing strategies available to a multihomed customer connected to multiple service providers
- Describe AS numbering strategies available to a multihomed customer connected to multiple service providers
- Describe the operation of AS number translation
- Describe how you can implement a typical backup setup between a multihomed customer and multiple service providers in a BGP environment
- Describe the use of BGP attributes to influence inbound link selection in customer networks multihomed to multiple service providers
- Describe how you can implement load sharing between a multihomed customer and multiple service providers in a BGP environment

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

 Copyright © 2003, Cisco Systems, Inc.

Outline

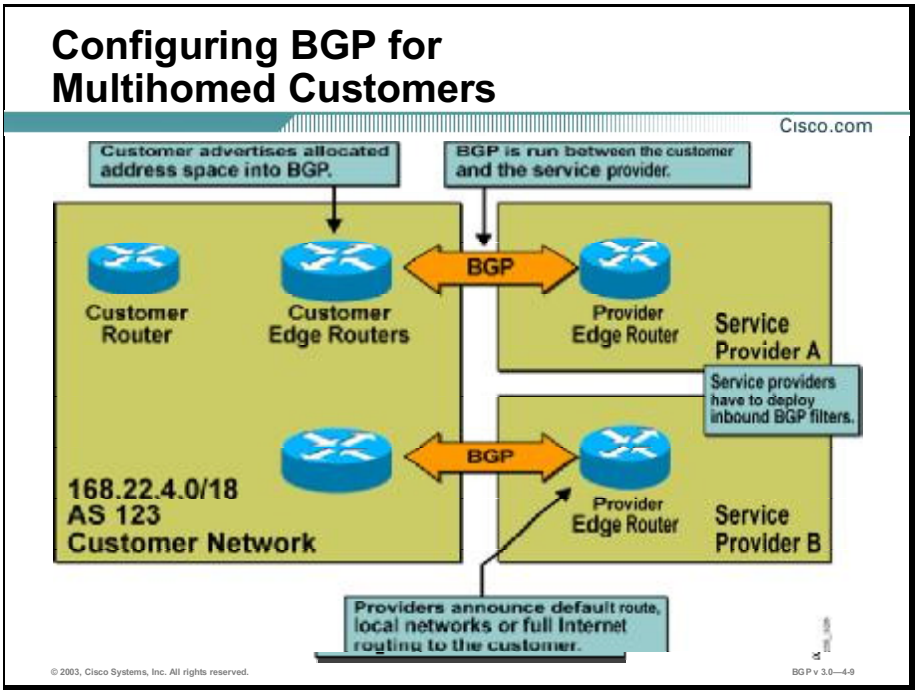
This lesson includes these topics:

- Overview
- Configuring BGP for Multihomed Customers
- Multihomed Customer Address Space Selection
- Multihomed Customer AS Number Selection
- AS Number Translation
- Primary/Backup Link Selection
- BGP Incoming Link Selection
- Load Sharing with Multiple Providers
- Summary
- Assessment (Quiz): Connecting a Multihomed Customer to Multiple Service Providers

Copyright © 2003, Cisco Systems, Inc.

Configuring BGP for Multihomed Customers

This topic describes the different characteristics of a BGP configuration used to establish routing between a multihomed customer and multiple service providers.


Configuring BGP for Multihomed Customers
Cisco.com
Customer advertises allocated address space into BGP.
BGP is run between the customer and the service provider.
Customer Router
Customer Edge Routers
BGP
Provider Edge Router
Service Provider A
Service providers have to deploy inbound BGP filters.
BGP
168.22.4.0/18
AS 123
Customer Network
Provider Edge Router
Service Provider B
Providers announce default route, local networks or full Internet routing to the customer.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-9

The highest level of resilience to network failures is achieved in network designs that connect the customer network to two different service provider networks. Customers use this option when the requirement for resilient Internet connectivity is very high. This requirement also involves duplication of equipment to make the customer network fully redundant.

BGP must be used between the customer and both service providers, because static routing will not work in this type of network. It is not enough to detect link failures or a failure in the remote router by link-level procedures. Failures that occur beyond the directly connected router must also be detected, and the only means of detecting these failures is by using a routing protocol. The only routing protocol suited for the Internet environment is BGP. Correctly configured, BGP takes care of rerouting in the following situations:

- Link failure between the customer network and the network of one of the ISPs.
- Edge router failure on either the customer or the ISP side.
- Link failure or router failure within the customer network causing the customer edge router to lose connectivity with the customer network core. This situation requires correct configuration of route advertisement as described in an earlier lesson.
- Link failure or router failure within the ISP network causing the ISP edge router to lose connectivity with the rest of the Internet.

 Copyright © 2003, Cisco Systems, Inc.

Multihomed customers have multiple permanent links to different ISPs. The links should terminate in different edge routers in the customer network. Otherwise, one of the major advantages, resilience to router failure, is lost.

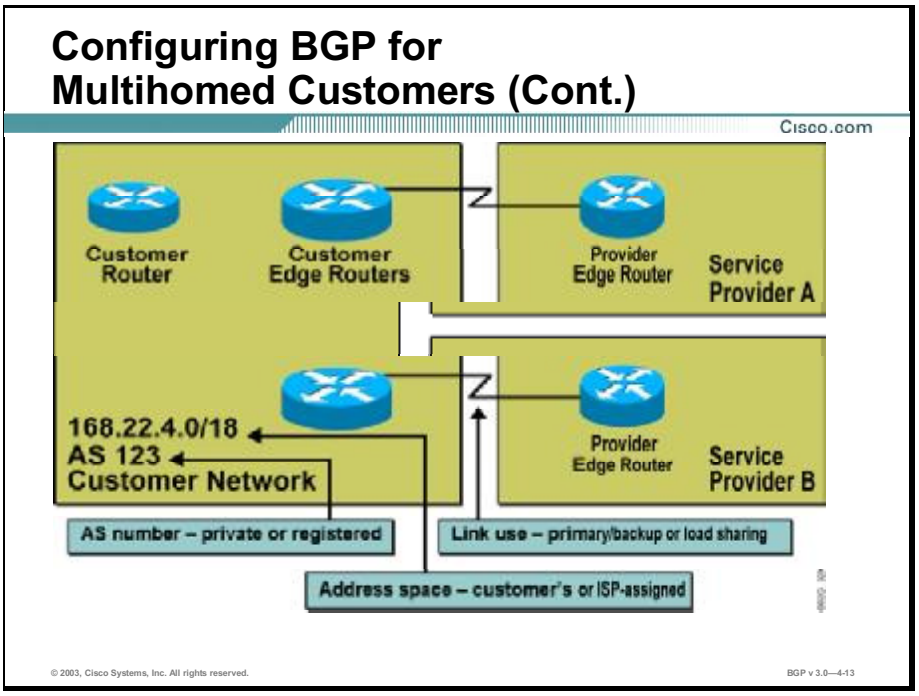
Multihomed customers should use BGP with both ISPs. The customer should advertise its address space to both providers. Route advertisement should be configured in both customer edge routers. The advertisement should be conditioned with the appropriate route policies by the edge router connectivity to the core of the customer network. This setup is analogous to what is configured when you are connecting a multihomed customer network to a single provider.

The customer should take care not to move any routing information between the two ISPs. It must use outgoing filters to prevent any route received from one of the ISPs to be propagated to the other. Otherwise, the customer network appears as a transit network between the two ISPs.

Both ISP must apply filters on the incoming BGP information from the customer to protect themselves and the rest of the Internet from errors in the BGP configuration of the customer. Each of the service providers must accept routes from the customer that indicate networks within the customer address space only. AS-path filter-lists should be implemented on the provider edge routers to allow incoming routes only if they have the correct AS-path attribute value. If the incoming filters on the ISP edge router accept customer routes, then the service provider should propagate those routes to the rest of the Internet.

Both ISPs must provide the customer with at least some BGP routes. Depending on customer requirements, the volume of BGP routes provided by the ISP could range anywhere from the default route only, to the full Internet routing table.

Copyright © 2003, Cisco Systems, Inc.



Before configuring the multihomed network, you need to consider the following questions:

- Should any of the links be used as primary and the others as backup?
- Should both links share the load?
- What address space is the customer using? (Is the customer address space provider-assigned [PA] or provider-independent [PI]?)
- What AS number is the customer using? (Is the customer using a public or a private AS number?)

Practice

Q1) Why should you apply outbound filters to a multihomed customer network?

A) to provide for maximum security at the customer site

B) to guarantee that return traffic has proper load balancing

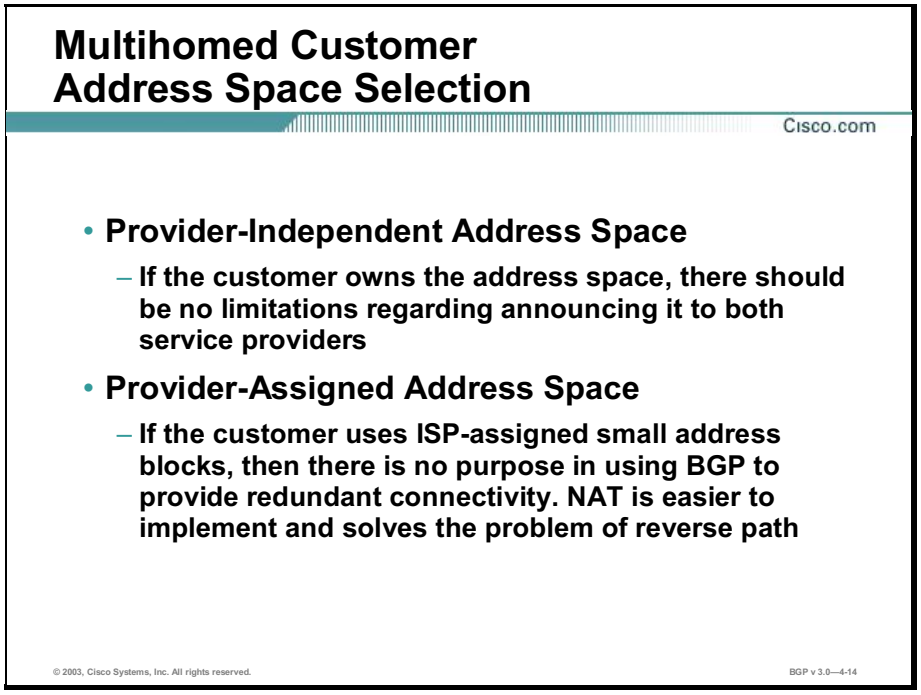
C) to prevent the customer network from becoming a transit AS

D) to ensure that outbound traffic has proper load balancing from the customer site

 Copyright © 2003, Cisco Systems, Inc.

Multihomed Customer Address Space Selection

This topic describes the different addressing strategies available to a multihomed customer connected to multiple service providers.


Multihomed Customer
Address Space Selection
Cisco.com
• Provider-Independent Address Space
– If the customer owns the address space, there should be no limitations regarding announcing it to both service providers
• Provider-Assigned Address Space
– If the customer uses ISP-assigned small address blocks, then there is no purpose in using BGP to provide redundant connectivity. NAT is easier to implement and solves the problem of reverse path
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-14

If the customer has its own address space, it should announce it to both service providers. Both providers are responsible for propagating the customer routes to the rest of the Internet without doing any summarization.

However, if the customer uses a small block of addresses assigned by one of the ISPs, an alternative design, not involving BGP, is to use two different PA address spaces and do Network Address Translation (NAT). With NAT, the router translates traffic going out over one of its connections to one of the PA addresses. If traffic goes out the other way, the addresses are translated to an address from the address space of the other provider.

Practice

Q1) Why is PI address space recommended for customers connected to multiple providers?

A) PI addressing facilitates easier implementation of backup link policies.

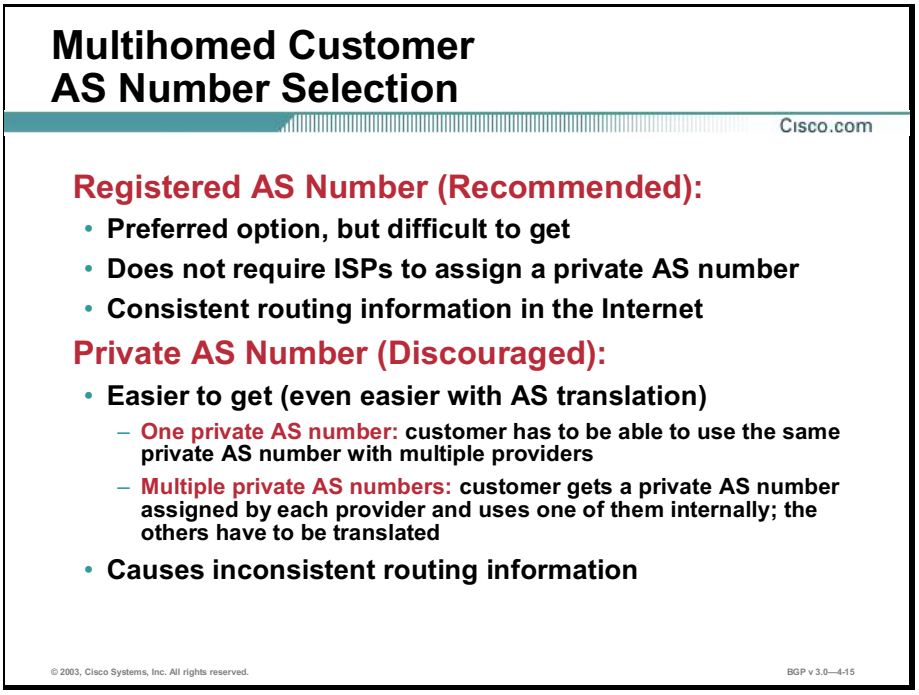
B) PI addressing facilitates easier implementation of load-sharing configurations.

C) PI addressing removes any limitations around advertising the address space.

D) PI addressing enables the use of automatic summarization at upstream ISPs.

Multihomed Customer AS Number Selection

This topic describes the different AS numbering strategies available to a multihomed customer connected to multiple service providers.


Multihomed Customer
AS Number Selection
Cisco.com
Registered AS Number (Recommended):
• Preferred option, but difficult to get
• Does not require ISPs to assign a private AS number
• Consistent routing information in the Internet
Private AS Number (Discouraged):
• Easier to get (even easier with AS translation)
– One private AS number: customer has to be able to use the same private AS number with multiple providers
– Multiple private AS numbers: customer gets a private AS number assigned by each provider and uses one of them internally; the others have to be translated
• Causes inconsistent routing information
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-15

The use of BGP requires an AS number. The preferred option is to use a registered, public AS number. However, registered AS numbers are assigned only to those who really need it because public AS numbers are a scarce resource. An ISP with BGP sessions to multiple ISPs must use a registered, public AS number. A customer connected to only one ISP does not require a public AS number. In that case, a private AS number in the range 64512-65535 is sufficient.

Whenever the customer has a public AS number assigned, there are no conflicts in the BGP setup, because the number is guaranteed to be unique within the Internet. Route announcements are made by both the customer and service provider without tampering with the AS path. As a result, consistent AS path information is propagated by the service provider to the rest of the Internet.

In those cases where the customer does not have a public AS number, it must use a private AS number. Because private AS numbers are not propagated to the Internet, several network administrators can, independently of each other, make this assignment. In this case, AS numbers are reused, conserving AS number space. A service provider normally assigns private AS numbers to its customers. This arrangement makes sure that unique private AS numbers are used among the customers of a single ISP.

 Copyright © 2003, Cisco Systems, Inc.

In the case where a customer is going to be multihomed and the private AS number already assigned from one of the ISPs comes in conflict with AS numbers assigned by the other ISP, renumbering of the customer AS should be considered by the customer. If the two service providers can reach a common agreement on which private AS number the multihomed customer should use, renumbering is a solution. If no common agreement can be made or if renumbering, for some reason, is not an option, AS translation must be configured on the customer network.

No router should ever propagate private AS numbers to the rest of the Internet. An ISP can keep track of which private AS numbers it has assigned to its customers and avoid reuse or conflicts within that scope. However, as soon as the scope is widened to include other ISPs, conflicts will happen. Each ISP, therefore, removes private AS numbers from the AS path before sending routes outside its own AS.

When the routes, with the private AS numbers removed, are propagated to the rest of the Internet, the AS path looks like the routes were originated within the public AS of the ISP. All information about the private AS lying behind the public AS is lost. In the case of a multihomed customer, the customer routes are, in the first step, propagated into each of the autonomous systems of its ISPs. In the next step, the routes have the private AS number removed as the routes are propagated to the rest of the Internet. Now the customer routes appear to be originating in the autonomous systems of both ISPs. To an outside observer, there is now an AS-path inconsistency because it looks like the same route belongs to different autonomous systems.

Practice

Q1) What are three AS number implementation options available to customers connecting to multiple providers? (Choose three.)

A) The customer can obtain a registered, public AS number that is advertised to all upstream providers.

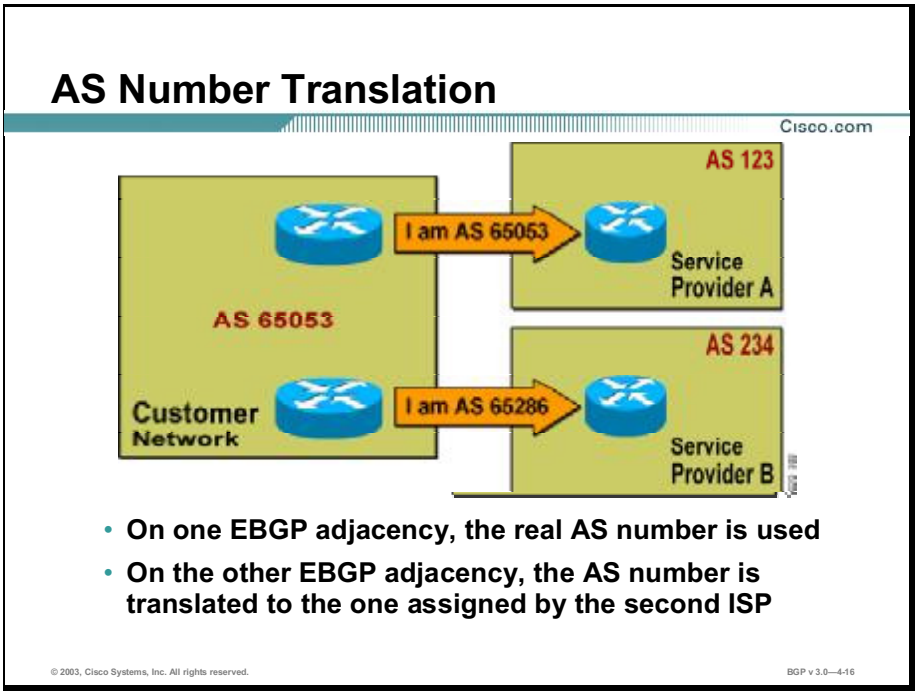
B) The customer can use a single private AS number as long as all upstream providers agree to support the same AS number.

C) The customer can use two different private AS numbers by translating one of the private addresses at the customer edge.

D) The customer can use two different PA AS numbers by configuring EBGP internally at the customer site.

AS Number Translation

This topic describes the operation of AS number translation.


AS Number Translation
Cisco.com
AS 123
I am AS 65053
Service Provider A
AS 65053
AS 234
Customer Network
I am AS 65286
Service Provider B
• On one EBGP adjacency, the real AS number is used
• On the other EBGP adjacency, the AS number is translated to the one assigned by the second ISP
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-16

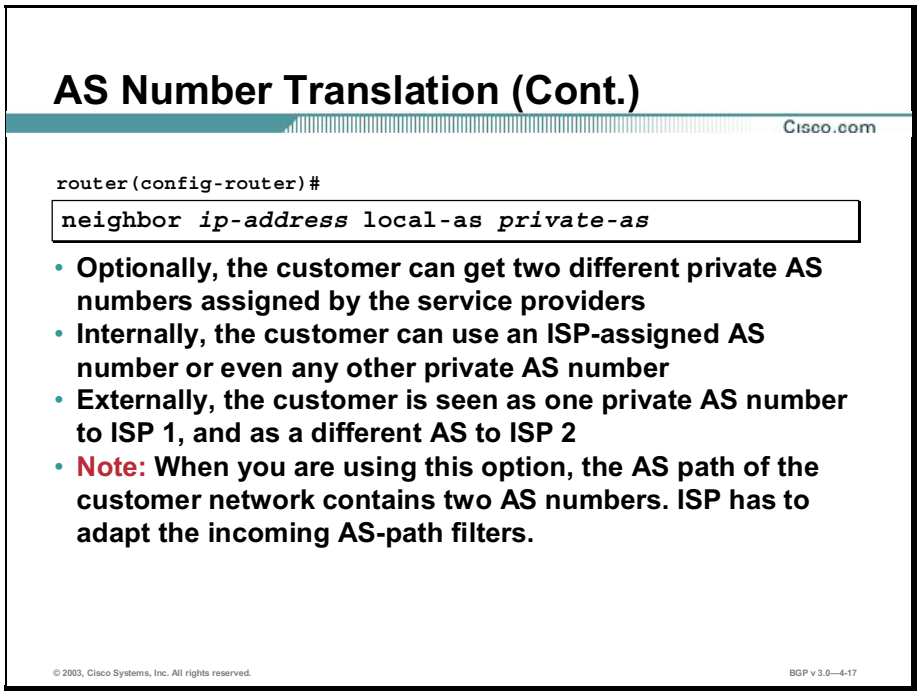
The figure shows a case where a customer is multihomed but forced to use two private AS numbers (for example, because of the scarcity of public AS numbers).

In the figure, service provider A has assigned the private AS number 65053 to the customer. Service provider B did not agree to use this private AS number when connecting to the customer. Instead, service provider B has assigned the private AS number 65286.

The customer now has two different private AS numbers: 65053 and 65286. The customer decides to use 65053 internally. All router BGP configuration lines have 65053 as the AS number. The customer uses AS number 65286 only when establishing the External Border Gateway Protocol (EBGP) session to AS 234.

In the example, service provider A (AS 123) has an EBGP session to the customer where the AS number 65053 is used at the customer end. Service provider B (AS 234) has an EBGP session to the customer where the AS number 65286 is used at the customer end. Translation between these two private AS numbers takes place in the customer edge router as part of the EBGP session to AS 234.

 Copyright © 2003, Cisco Systems, Inc.


AS Number Translation (Cont.)
Cisco.com
router(config-router)#
neighbor ip-address local-as private-as
• Optionally, the customer can get two different private AS numbers assigned by the service providers
• Internally, the customer can use an ISP-assigned AS number or even any other private AS number
• Externally, the customer is seen as one private AS number to ISP 1, and as a different AS to ISP 2
• Note: When you are using this option, the AS path of the customer network contains two AS numbers. ISP has to adapt the incoming AS-path filters.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-17

The **neighbor** *ip-address* **local-as** *private-as* router configuration command is used to indicate the AS number that the local router uses as its local AS number in the BGP Open message. The remote router is assumed to have an EBGP session to the indicated local AS.

Internally, the customer network uses another private AS number. When routes are sent to the neighbor, the internal AS number is automatically prepended in the AS path first, and then the specified local AS number is prepended as well. As a consequence, the ISP receives the routes with an AS path with both AS numbers in it. The ISP has to adapt its incoming filter-lists as a result of this situation.

Note	Some service providers might be unwilling to change their AS-path input filters, leaving the customer no other option than using a public AS number or connecting to a single ISP with a private AS number.

Practice

Q1) What are two issues that arise when you are using AS number translation? (Choose two.)

A) The upstream provider must not filter routes based on a single AS.

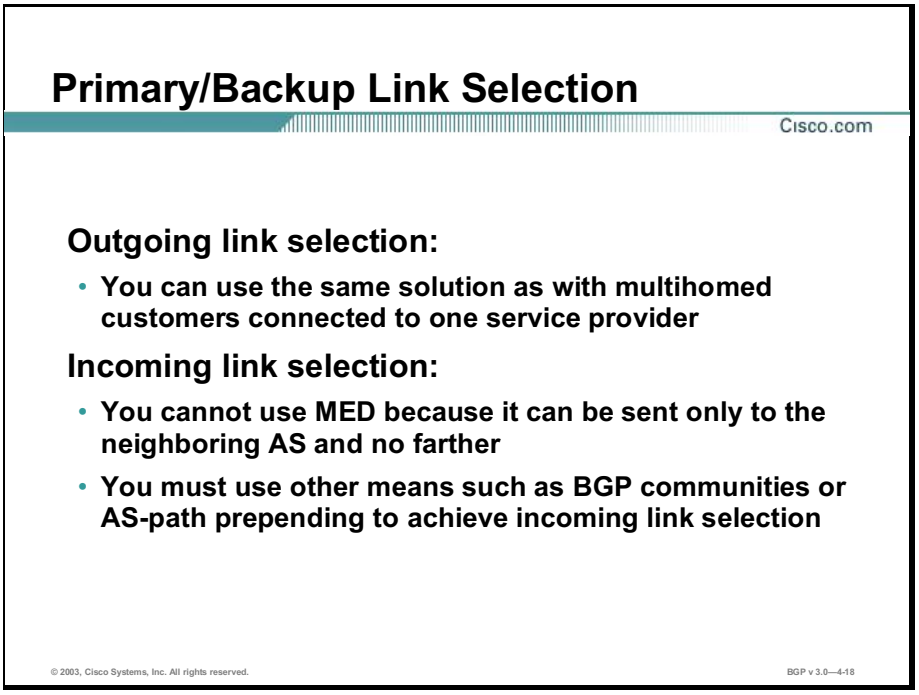
B) AS-path prepending is not supported when you are using AS number translation.

C) AS number translation causes two AS numbers to be prepended to the AS path.

D) You can use only private AS numbers with AS number translation.

Primary/Backup Link Selection

This topic describes how you can implement a typical backup setup between a multihomed customer and multiple service providers in a BGP environment.


Primary/Backup Link Selection
Cisco.com
Outgoing link selection:
• You can use the same solution as with multihomed customers connected to one service provider
Incoming link selection:
• You cannot use MED because it can be sent only to the neighboring AS and no farther
• You must use other means such as BGP communities or AS-path prepending to achieve incoming link selection
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-18

When using BGP on multiple links between a customer and several service provider networks, the customer is solely responsible for controlling the use of the links between them for outgoing traffic. The customer chooses whether to use these links in a primary/backup or a load-sharing configuration.

If one link is primary and the other used for backup purposes only, the customer can use the local preference attribute in the configuration to direct all outgoing traffic over the primary link. This configuration is no different than the configuration used for customers with multiple connections running BGP to a single service provider.

Controlling the load distribution of incoming traffic over multiple links is more difficult in the multihomed scenario when links to multiple service providers are used. You cannot use MED when the customer connects to multiple providers because the updates are sent to two different autonomous systems. Recall that MED is used only when comparing routes received from a single directly connected AS over two parallel links. Therefore, route selection decisions will most likely use the AS-path attribute and prefer the route with the shortest AS-path length.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) In what two ways is the primary/backup design different from the one used for multihomed customers connected to a single ISP? (Choose two.)

A) You cannot accomplish incoming route selection using the MED attribute.

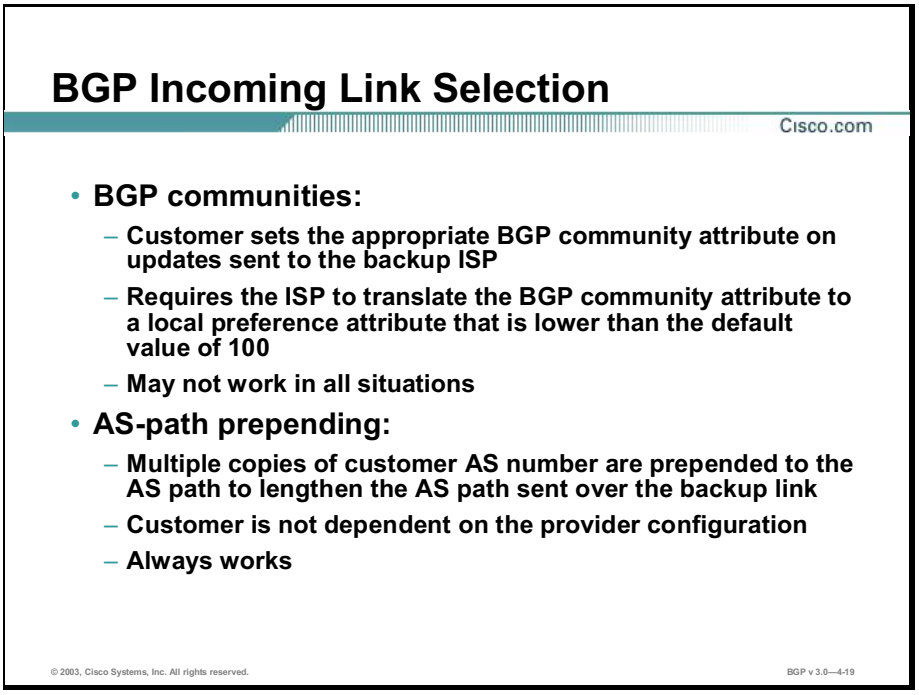
B) You can accomplish outgoing route selection only by using weights.

C) The customer should use local preference to direct traffic to the correct outgoing link.

D) BGP communities and AS-path prepending are used to influence incoming route selection.

BGP Incoming Link Selection

This topic describes the use of BGP attributes to influence inbound link selection in customer networks multihomed to multiple service providers.


BGP Incoming Link Selection
Cisco.com
• BGP communities:
– Customer sets the appropriate BGP community attribute on updates sent to the backup ISP
– Requires the ISP to translate the BGP community attribute to a local preference attribute that is lower than the default value of 100
– May not work in all situations
• AS-path prepending:
– Multiple copies of customer AS number are prepended to the AS path to lengthen the AS path sent over the backup link
– Customer is not dependent on the provider configuration
– Always works
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-19

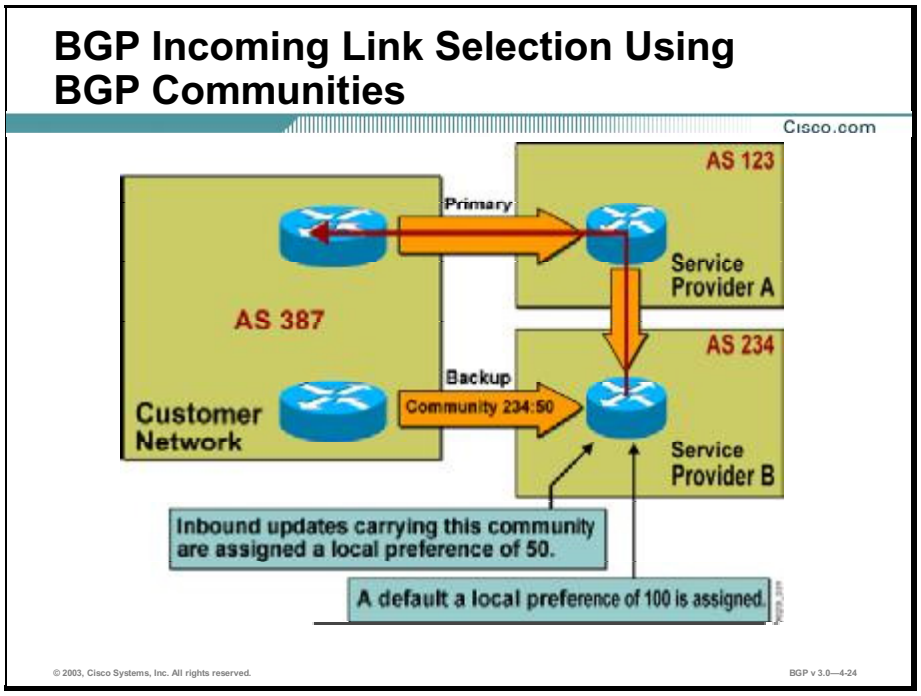
In order to remove incoming traffic from the backup link, the customer must influence route selection in the backup AS. The backup ISP must be forced to prefer the primary path to reach the customer network, although this choice means selecting a route with a longer AS path.

One way to influence route selection is to use local preference in the network of the backup ISP. Using local preference creates an administrative scalability issue if each customer requires its use, because the ISP must maintain the configuration.

One scalable way of setting local preference in an ISP network is to use communities. The customer sets a well-known community value on the routes sent to the backup ISP. The ISP recognizes the community and sets the local preference for these routes. This solution is available only if the ISP has implemented and announced the use of communities. If communities and a local preference setting are used, route selection occurs only if there are alternative routes to compare.

Another way of influencing route selection in the backup ISP is to do AS-path prepending before sending the advertisement to the backup ISP. When the customer sends routes over the backup link, multiple copies of its own AS number are prepended to the AS path of each route. The backup ISP receives the routes and makes normal route selection decisions. No special weight or local preference settings are used; the BGP route selection is based exclusively on the AS-path length. No special configuration is required in the service provider network.

 Copyright © 2003, Cisco Systems, Inc.



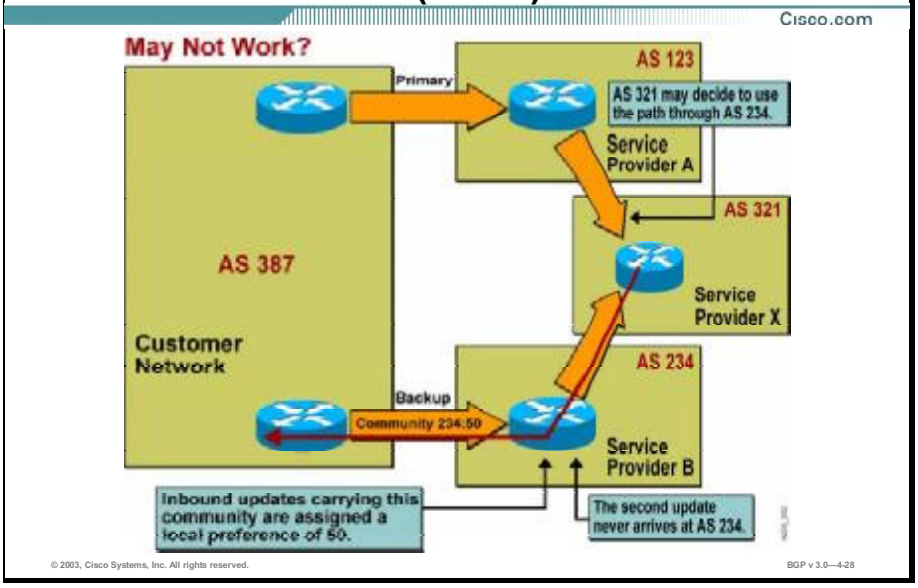
The backup service provider B (AS 234) has defined the meaning of community 234:50. When AS 234 receives routes with this community, the local preference is set to 50.

The customer AS 387 is advertising the route over the primary link without any communities. It is received by AS 123 and propagated to AS 234. When AS 234 receives the route via AS 123, there is no community set. AS 234 therefore assigns the default local preference value of 100.

The customer is also advertising the route over the backup link. However, in this case, the route has the community 234:50 set. When AS 234 receives this route, it recognizes the community, and the local preference value is set to 50.

Route selection is now performed in AS 234. The route received via AS 123 is preferred based on the local preference values.

BGP Incoming Link Selection Using BGP Communities (Cont.)



Although the use of communities is correctly configured, the desired load distribution may not always be achieved. As this example shows, AS 234 does not always receive the primary route although nothing is wrong in the network.

The customer AS 387 sends routes with community 234:50 over the backup link to AS 234. AS 234 receives the routes and sets the local preference to 50. If AS 234 over some period of time selects the directly connected path to AS387as the best, it propagates the route to AS 321. As the route is propagated over the EBGP session between AS 234 and AS 321, the local preference value used within AS 234 is lost.

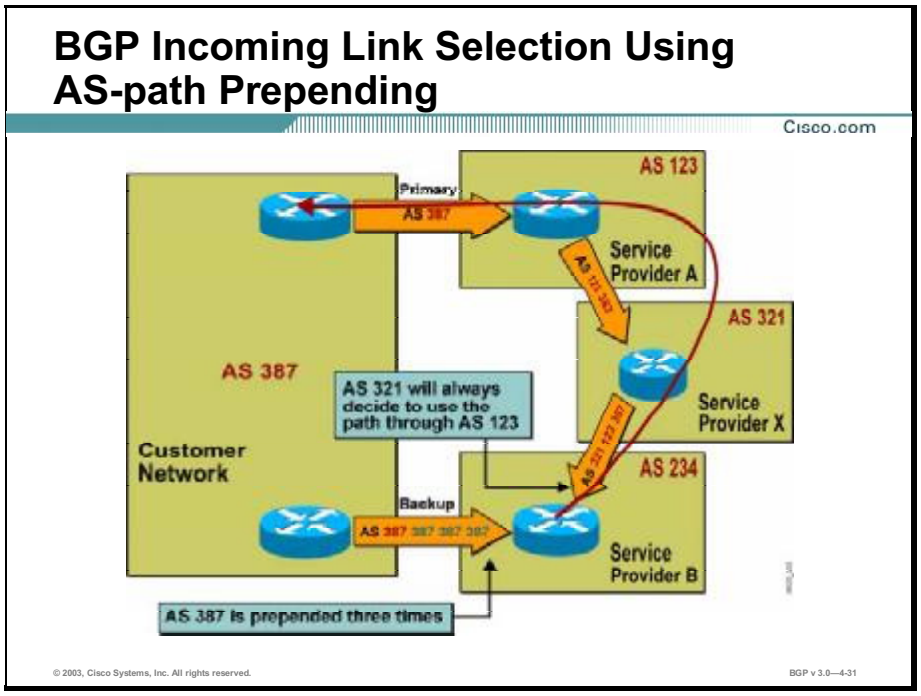
AS 321 does not have any use for the community 234:50 because this community is defined and implemented only within AS 234. Potentially, the community value can also be stripped off during BGP route propagation.

Customer AS 387 also sends the routes over the primary link to AS 123. The routes are propagated to AS 321, which now sees two alternative routes to the destination networks within AS 387. Neither weight nor local preference is used by the routers in AS 321as criteria for reaching AS 387 for those routes within AS 321. Both alternatives have equal AS-path length.

The route selection decision that will be made in AS 321 is hard to predict, but the outcome definitely influences the route selection decision made in AS 234. If AS 321 prefers the route to the customer network via AS 234 for any reason, then the second-best alternative via AS 123 and the primary link is never propagated to AS 234.

In this case, AS 234 never sees the primary path and has to stick to the backup link and announce the route to AS 321. The network has reached a steady state when the traffic uses the backup link although the primary link is available.

 Copyright © 2003, Cisco Systems, Inc.



In this example, the customer AS 387 is performing AS-path prepending on the backup link. Three copies of the customer AS number (387) are prepended to the AS path. As the route goes out over the EBGP session, BGP prepends the local AS number to the AS-path attribute. AS 234 receives routes from AS 387 over the backup link with an AS-path length of four (the original AS 387 plus three prepended copies that the customer edge router applied to the AS-path attribute).

The customer advertises networks without AS-path prepending over the primary link. AS 123 receives routes with an AS-path length of one and propagates these routes to AS 321, which then receives them with an AS-path length of two.

If, for a short period of time, AS 321 received the customer routes via AS 234, the AS-path length of those routes would have been five. In that case, AS 321 selects the route from AS 123 as the best and propagates it to AS 234.

AS 234 now sees both alternatives. The customer routes received directly from the customer have an AS-path length of four. The routes received via AS 321 have an AS-path length of three. Because no weight or local preference is configured in this example, AS 234 selects the route via AS 321 as the best.

The desired result, to have all traffic enter the customer network via the primary link, is now achieved.

Note If the backup ISP is implementing incoming AS-path filters for this customer with the length of the AS path equal to one, the ISP has to change the configuration of the AS-path filter for the customer. The ISP can either create a new filter, allowing multiple copies of the customer AS number only for this customer, or create a common filter for all customers belonging to one peer group, using regular expression variables.

Practice

Q1) What are three benefits of using AS-path prepending instead of BGP communities when you are influencing incoming path selection? (Choose three.)

A) The customer is not dependent upon the service provider for configuration changes and maintenance.

B) Using BGP communities to set local preference may not work in all network scenarios.

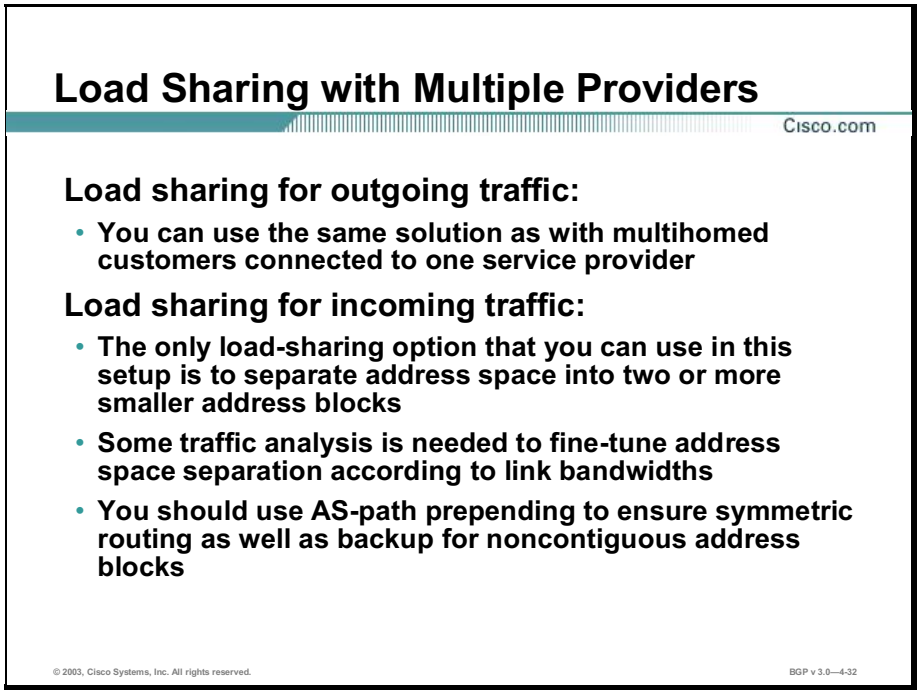
C) AS-path prepending requires only that the backup provider support multiple instances of the same AS number in the AS path to function correctly.

D) Using BGP communities to set local preference cannot be scaled in a service provider network.

 Copyright © 2003, Cisco Systems, Inc.

Load Sharing with Multiple Providers

This topic describes how you can implement load sharing between a multihomed customer and multiple service providers in a BGP environment.


Load Sharing with Multiple Providers
Cisco.com
Load sharing for outgoing traffic:
• You can use the same solution as with multihomed customers connected to one service provider
Load sharing for incoming traffic:
• The only load-sharing option that you can use in this setup is to separate address space into two or more smaller address blocks
• Some traffic analysis is needed to fine-tune address space separation according to link bandwidths
• You should use AS-path prepending to ensure symmetric routing as well as backup for noncontiguous address blocks
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-32

Load sharing over links to two different ISPs can be compared to doing load sharing over two parallel links to a single ISP. The only difference is that there is only one option available to control incoming traffic. Controlling load distribution of the outgoing traffic is configured in exactly the same way as when a multihomed customer connects to a single service provider.

The customer can control the load distribution of incoming traffic based on traffic destination. The customer divides its address space into several announcements. One announcement is sent to one of the ISPs. Another announcement is sent to the other ISP. For backup purposes, the customer announces the entire address space to both ISPs. The ISPs now use the most explicit route rule, and as long as both links are up, traffic with destinations within one part of the customer address space is routed over one of the links and traffic to the other part is routed over the other link.

It is very difficult to predict the volume of traffic that will be directed to one part of the customer address space and the volume that will be directed to the other part. You should monitor the results of changing route updates by watching the load on the links before and after implementing the change. If the load distribution is not satisfactory, you can further modify the division of the address space. You must then check the load on the links again and further fine-tune the configuration.

A customer may decide to use both the division of address space into several advertisements and AS-path prepending together. Some part of the customer's address space may be advertised by the customer network with a longer AS path over one of the links to fine-tune the load. Also, there may be cases where there are noncontiguous subnets that cannot be divided because the prefixes would be too long. These subnets are evenly distributed between the links in a primary/backup configuration.

Practice

Q1) In what two ways is the load-sharing design different from the one used for multihomed customers connected to a single ISP? (Choose two.)

A) There is no difference in the design for outgoing traffic.

B) There is no difference in the design for incoming traffic.

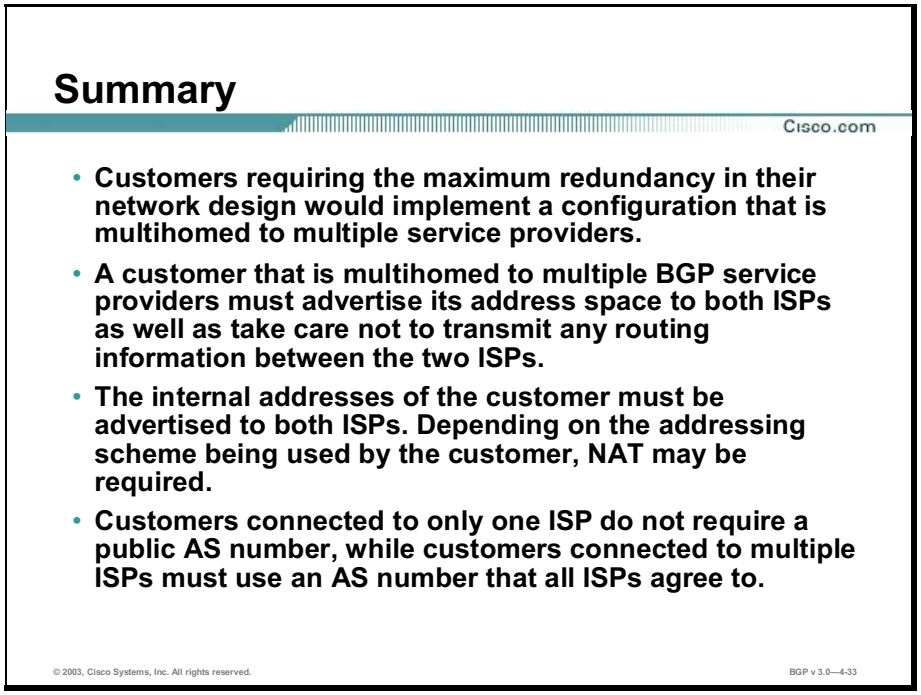
C) You can use AS-path prepending.

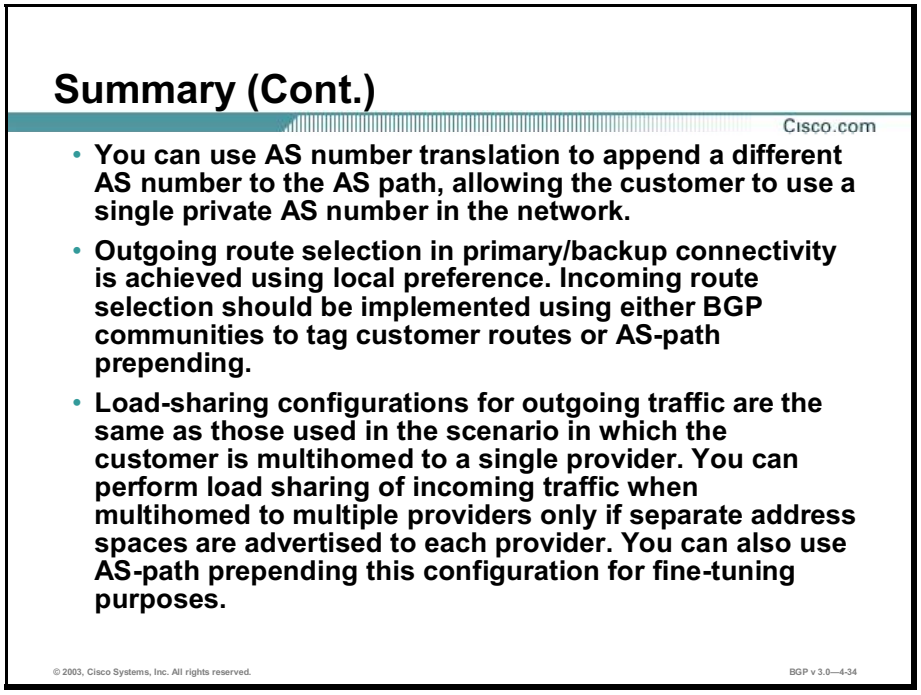
D) The weight, MED and local preference attributes all have to be altered.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Customers requiring the maximum redundancy in their network design would implement a configuration that is multihomed to multiple service providers.
• A customer that is multihomed to multiple BGP service providers must advertise its address space to both ISPs as well as take care not to transmit any routing information between the two ISPs.
• The internal addresses of the customer must be advertised to both ISPs. Depending on the addressing scheme being used by the customer, NAT may be required.
• Customers connected to only one ISP do not require a public AS number, while customers connected to multiple ISPs must use an AS number that all ISPs agree to.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-33


Summary (Cont.)
Cisco.com
• You can use AS number translation to append a different AS number to the AS path, allowing the customer to use a single private AS number in the network.
• Outgoing route selection in primary/backup connectivity is achieved using local preference. Incoming route selection should be implemented using either BGP communities to tag customer routes or AS-path prepending.
• Load-sharing configurations for outgoing traffic are the same as those used in the scenario in which the customer is multihomed to a single provider. You can perform load sharing of incoming traffic when multihomed to multiple providers only if separate address spaces are advertised to each provider. You can also use AS-path prepending this configuration for fine-tuning purposes.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v 3.0—4-34

Next Steps

After completing this lesson, go to:

- BGP Transit Autonomous System module

References

For additional information, refer to these resources:

- For more information on the BGP local AS feature, refer to “Configuring the BGP Local-AS Feature” at the following URL: http://www.cisco.com/warp/public/459/39.html
- For more information on load sharing, refer to “Sample Configurations for Load Sharing with BGP in Single and Multihomed Environments” at the following URL: http://www.cisco.com/warp/public/459/40.html

 Copyright © 2003, Cisco Systems, Inc.

Quiz: Connecting a Multihomed Customer to Multiple Service Providers

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe BGP configuration characteristics used to establish routing between a multihomed customer and multiple service providers
- Describe addressing strategies available to a multihomed customer connected to multiple service providers
- Describe AS numbering strategies available to a multihomed customer connected to multiple service providers
- Describe the operation of AS number translation
- Describe how you can implement a typical backup setup between a multihomed customer and multiple service providers in a BGP environment
- Describe the use of BGP attributes to influence inbound link selection in customer networks multihomed to multiple service providers
- Describe how you can implement load sharing between a multihomed customer and multiple service providers in a BGP environment

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Copyright © 2003, Cisco Systems, Inc.

Q1) A multihomed customer is using AS number 65500 internally. The customer is connected to two different providers. Provider 1 (in AS 222) has assigned the customer an AS number of AS 65101. Provider 2 (in AS 333) has assigned the customer an AS number of AS 65201. Given that the customer will use AS number translation for its internal AS, what is the AS-path attribute, attached to routes originated in the customer network, when displayed on a router in the network of Provider 2?

A) 65550 i

B) 65201 i

C) 65201 65550 i

D) 333 65201 i

Q2) What three methods can you use to provide load sharing over network links between a multihomed customer and multiple providers? (Choose three.)

A) advertising of split addressing space to the provider

B) configuring of multiple static routes pointing to the provider

C) use of the BGP **maximum-paths** command to perform load sharing over parallel links

D) AS-path prepending to fine-tune the load-sharing configuration

Q3) What are three BGP configuration characteristics of a multihomed customer connected to multiple providers? (Choose three.)

A) The customer announces assigned addressing to its providers through BGP.

B) The customer announces a default route to its network through BGP.

C) The provider announces a default route, local routes, or full Internet routing to the customer via BGP.

D) The customer configures outbound filters to prevent its network from becoming a transit area.

 Copyright © 2003, Cisco Systems, Inc.

Q4) A multihomed customer is using AS number 1024 and is connected to two different providers (Provider 1: AS 222 and Provider 2: AS 333). The customer has configured MED to ensure a proper return path so that Provider 1 is the primary provider and Provider 2 is the backup provider. Unfortunately, return traffic continues to use the backup link. What is a possible cause of this problem?

A) The backup provider is ignoring the MED attribute on received routes.

B) The MED attribute cannot be sent to the backup provider because it is local to AS 1024 only.

C) The customer has not set the proper BGP communities to allow the primary and backup providers to correctly set the MED attribute.

D) MED cannot be used in this scenario, because it will not be advertised to providers upstream of Provider 2.

Q5) What are three important considerations for customers wishing to connect to multiple providers? (Choose three.)

A) The customer has to consider whether to use PA or PI address space.

B) The customer has to decide whether to use static routes or BGP to connect to upstream providers.

C) The customer has to decide whether to use a public AS number or a private AS number scheme.

D) The customer has to decide whether to perform load sharing or use a primary/backup implementation over redundant links.

Q6) Which AS number selection is the best possible choice for a customer multihomed to multiple providers?

A) a single public AS number

B) a single private AS number

C) two private AS numbers used in conjunction with AS number translation

D) multiple private AS numbers, one used internally by the customer and the others used in conjunction with AS number translation for each provider

Q7) Given the following router command output, what two methods have been configured to influence return traffic in a primary/backup link for this multihomed customer? (Choose two.)

```
Provider# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.10.20.0/24    192.168.63.3          2000             0 100 100
100 i
*>                  192.168.64.4                           0 300 100
i
*> 30.30.30.0/24    192.168.63.3             0             0 300 100
i
*> 40.40.40.0/24    192.168.64.4             0             0 100 i
```

A) MED

B) local preference

C) split address advertisement

D) AS-path prepending

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

5

BGP Transit Autonomous Systems

Overview

This module is one of the focal points of the Border Gateway Protocol (BGP) curriculum: a discussion of BGP issues in a transit autonomous system (AS). The module covers basic BGP issues in transit autonomous systems, ranging from synchronization between an Interior Gateway Protocol (IGP) and BGP to Internal Border Gateway Protocol (IBGP) full-mesh and next-hop requirements.

Upon completing this module, you will be able to:

- Describe the function of a transit AS and the need for IBGP
- Describe the interaction between EBGP and IBGP in relation to relevant BGP attributes, given a transit AS
- Describe the function of an IGP in forwarding packets through an AS
- Successfully configure an AS to act as a transit backbone, given a BGP network scenario
- Verify proper operation and perform the steps necessary to correct basic IBGP configuration errors, given a configured BGP transit network

Outline

The module contains these lessons:

- Transit Autonomous System Functions
- IBGP and EBGP Interaction in a Transit Autonomous System
- Packet Forwarding in Transit Autonomous Systems
- Configuring a Transit Autonomous System
- Monitoring and Troubleshooting IBGP in Transit Autonomous Systems

 Copyright © 2003, Cisco Systems, Inc.

Transit Autonomous System Functions

Overview

The topology of the Internet can be viewed as a series of connections between stub networks, multihomed networks, and transit autonomous systems. A multihomed autonomous system (AS) containing more than one connection to the outside world and allowing traffic not originating in that AS to travel through it is a transit AS. This lesson introduces the concept of the multihomed transit AS and how Border Gateway Protocol (BGP) exchanges routing information inside the AS and between neighboring autonomous systems. It also explains the requirement for Internal Border Gateway Protocol (IBGP) within the multihomed transit AS.

Importance

All transit autonomous systems are required to carry traffic originating from and/or destined to locations outside of that AS. In order for the transit AS to achieve this, a degree of interaction and coordination between BGP and the Interior Gateway Protocol (IGP) used by that particular AS is necessary. Such a configuration requires special care to ensure consistency of routing information throughout the AS.

Objectives

Upon completing this lesson, you will be able to:

- List the functions of a transit AS
- Describe external route propagation between autonomous systems in a BGP network
- Describe internal route propagation within a BGP AS
- Explain how transiting packets are forwarded inside a transit AS
- Explain the need for deploying IBGP on all core routers

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

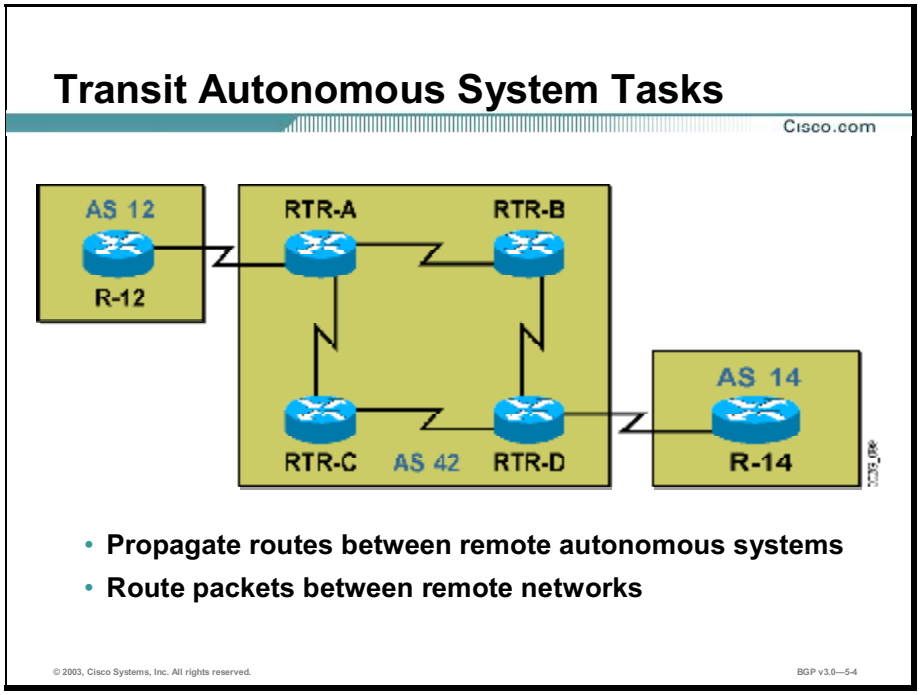
This lesson includes these topics:

- Overview
- Transit Autonomous System Tasks
- External Route Propagation
- Internal Route Propagation
- Packet Forwarding in an Autonomous System
- Core Router IBGP Requirements in a Transit Autonomous System
- Summary
- Assessment (Quiz): Transit Autonomous System Functions

 Copyright © 2003, Cisco Systems, Inc.

Transit Autonomous System Tasks

This topic lists the functions of a transit AS.



Routers in a transit AS have to perform two tasks:

- Receive routing information updates about reachable networks from neighboring autonomous systems, propagate the information through their own AS, and send it to other neighboring autonomous systems.
- Forward IP packets received from a neighboring AS through their own AS to a downstream neighboring AS. The routers in the transit AS perform this task using the routing information received in the first task.

Practice

Q1) What are the two main functions of a transit AS? (Choose two.)

A) to propagate routes between remote autonomous systems

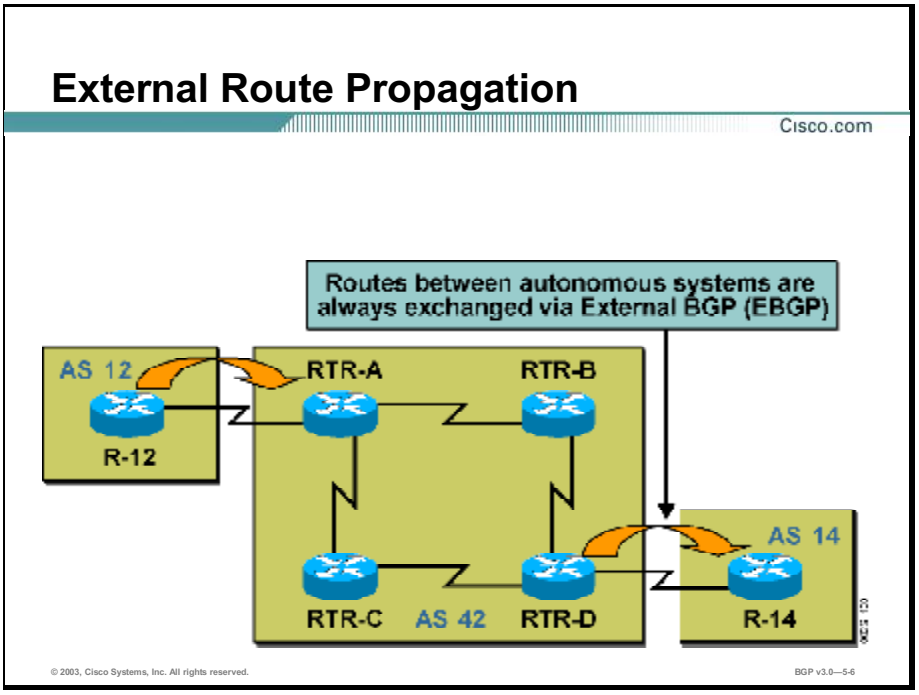
B) to filter noncustomer routes from transiting the AS

C) to route packets between remote networks

D) to connect customer networks to Internet service providers

External Route Propagation

This topic describes external route propagation between autonomous systems in a BGP network.



Two autonomous systems usually exchange routing information about reachable networks using BGP. There is currently no alternative routing protocol that has the scalability and security characteristics of BGP.

In the example here, the BGP session between R-12 and RTR-A is called an External Border Gateway Protocol (EBGP) session because R-12 and RTR-A are in different autonomous systems.

BGP routing information updates consist of the network address, subnet mask, and any number of BGP attributes. No other routing protocol provides the same richness of route attributes as BGP. Translating BGP route attribute information into any other protocol would likely cause a loss of information. Therefore, the EBGP information that RTR-A receives is not translated; it is just forwarded to other BGP-speaking routers (RTR-D in the figure) within the AS.

Likewise, RTR-D has BGP information and can propagate it to R-14 in AS 14 over the EBGP session.

EBGP sessions are, in general, established between directly connected neighbors. BGP-speaking routers thus need no additional routing information in order to establish the session.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) How are external routes exchanged between autonomous systems?

A) through EBGP

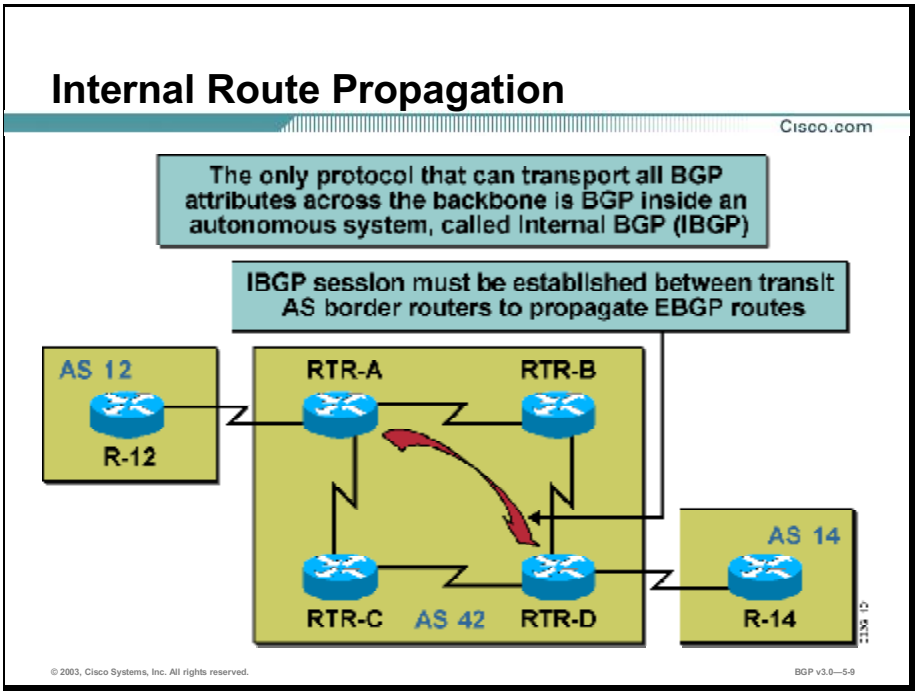
B) through IBGP

C) through route redistribution into an IGP

D) through local advertisement of routes at the edge router

Internal Route Propagation

This topic describes internal route propagation within a BGP AS.


Internal Route Propagation
Cisco.com
The only protocol that can transport all BGP attributes across the backbone is BGP inside an autonomous system, called Internal BGP (IBGP)
IBGP session must be established between transit AS border routers to propagate EBGP routes
AS 12
R-12
RTR-A
RTR-B
RTR-C
AS 42
RTR-D
AS 14
R-14
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-9

In this example, the BGP session between RTR-A and RTR-D, which are both in the same AS, is an IBGP session.

IBGP sessions are, in general, established between distant routers in the same AS. These routers need additional routing information in order to establish the session, because there is no requirement that IBGP neighbors be directly connected. This information typically comes from the IGP, which is running within the AS independently of BGP.

Practice

Q1) How are BGP routes propagated across an AS?

A) through EBGP

B) through IBGP

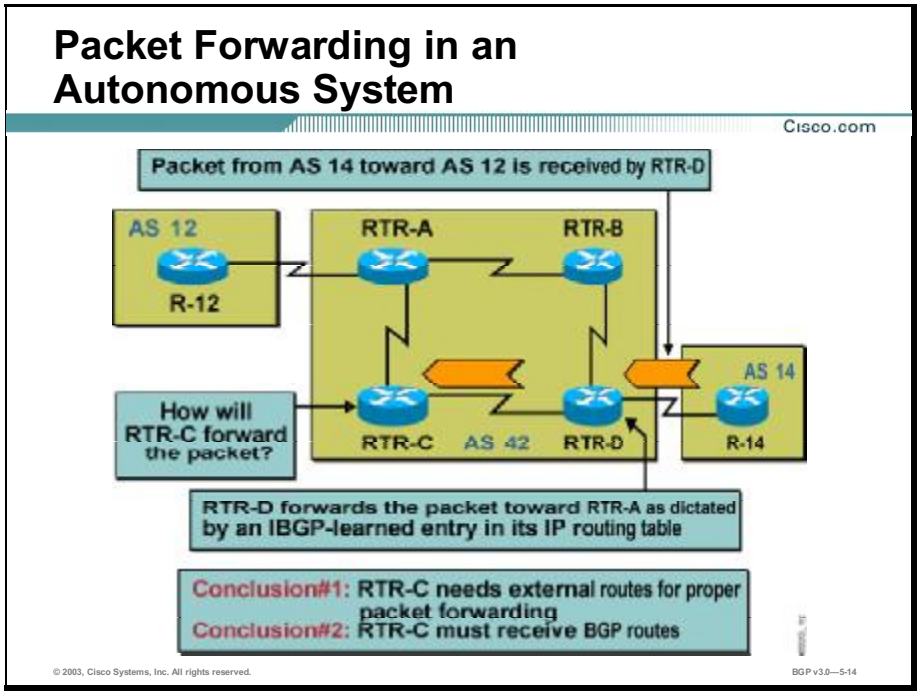
C) through route redistribution into an IGP

D) through local advertisement of routes at the edge router

 Copyright © 2003, Cisco Systems, Inc.

Packet Forwarding in an Autonomous System

This topic explains how transiting packets are forwarded inside a transit AS.


Packet Forwarding in an Autonomous System
Cisco.com
Packet from AS 14 toward AS 12 is received by RTR-D
AS 12
R-12
RTR-A
RTR-B
AS 14
How will RTR-C forward the packet?
RTR-C
AS 42
RTR-D
R-14
RTR-D forwards the packet toward RTR-A as dictated by an IBGP-learned entry in its IP routing table
Conclusion#1: RTR-C needs external routes for proper packet forwarding
Conclusion#2: RTR-C must receive BGP routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-14

In this example, after AS 14 has received the routing information about reachable networks inside AS 12, IP packets can start to flow (in the figure, from AS 14 toward AS 12). R-14, the egress router in AS 14, forwards IP packets with destinations in AS 12 toward RTR-D, according to information received through EBGP.

RTR-D now uses the IBGP information received from RTR-A and forwards the packets in the direction of RTR-A, which in this case means via RTR-C.

When the IP packets reach RTR-C, the router checks its routing table for a matching entry, but it fails to find one. The packet is dropped because the destination network is unreachable from the perspective of RTR-C.

This is, of course, an unacceptable situation. To prevent dropped packets due to unreachable networks, RTR-C must also have routing information about the networks reachable inside AS 12. The same information that RTR-D received from RTR-A over the IBGP session must be propagated to RTR-C.

Note	RTR-B has the same network reachability requirements as RTR-C, because RTR-D could forward the packets via RTR-B as well as via RTR-C.

Practice

Q1) Why do you have to run BGP on all core routers in a transit AS?

A) to eliminate the possibility of routing loops

B) to optimize the routing across the AS

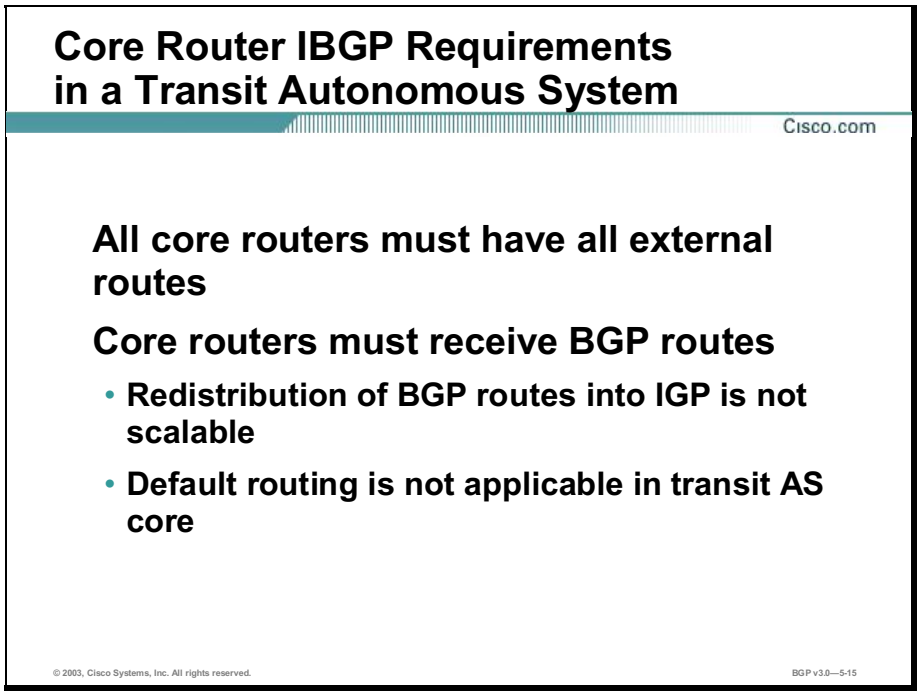
C) to be able to forward packets to all external destinations

D) to ensure that only one exit point exists for the transit backbone

 Copyright © 2003, Cisco Systems, Inc.

Core Router IBGP Requirements in a Transit Autonomous System

This topic explains the need for deploying IBGP on all core routers.


Core Router IBGP Requirements in a Transit Autonomous System
Cisco.com
All core routers must have all external routes
Core routers must receive BGP routes
• Redistribution of BGP routes into IGP is not scalable
• Default routing is not applicable in transit AS core
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-15

Within a transit AS, all routers, which are in a theoretical transit path between external destinations, should have information about all external routes received from any neighboring AS. If a single router on a transit path does not have this information, there is always a possibility that an IP packet received from a neighboring AS will not be able to be forwarded by that router through the transit AS. The router lacking routing information about the final destination of the IP packet drops it into what effectively becomes a black hole.

The only feasible way for the router to distribute all external routing information is by using IBGP. Redistribution of the EBGP routes into an IGP is not viable because no IGP can carry the volume of information that BGP currently carries in the Internet.

Note The risk of losing information when redistributing the EBGP routes into an IGP is not the reason why BGP is used to update intermediate routers in the transit path instead of an IGP. Redistribution into an IGP is not used because of the scalability issues that would arise from doing so.

Default routing or a gateway of last resort cannot be used by routers within the transit path when transit services are provided to other autonomous systems. If some routes were to be filtered out, and the default route used instead, full routing flexibility would be lost. The transit AS would not be able to forward packets to all destinations at all times. In fact, routing loops and black holes might be easily introduced.

Practice

Q1) Why is the redistribution of BGP routes into IGP not advisable?

A) Redistributing full Internet routing into any IGP would result in the loss of the BGP attributes needed to ensure optimal routing within an AS.

B) Redistributing full Internet routing is not possible, because BGP policies cannot be enforced by IGPs.

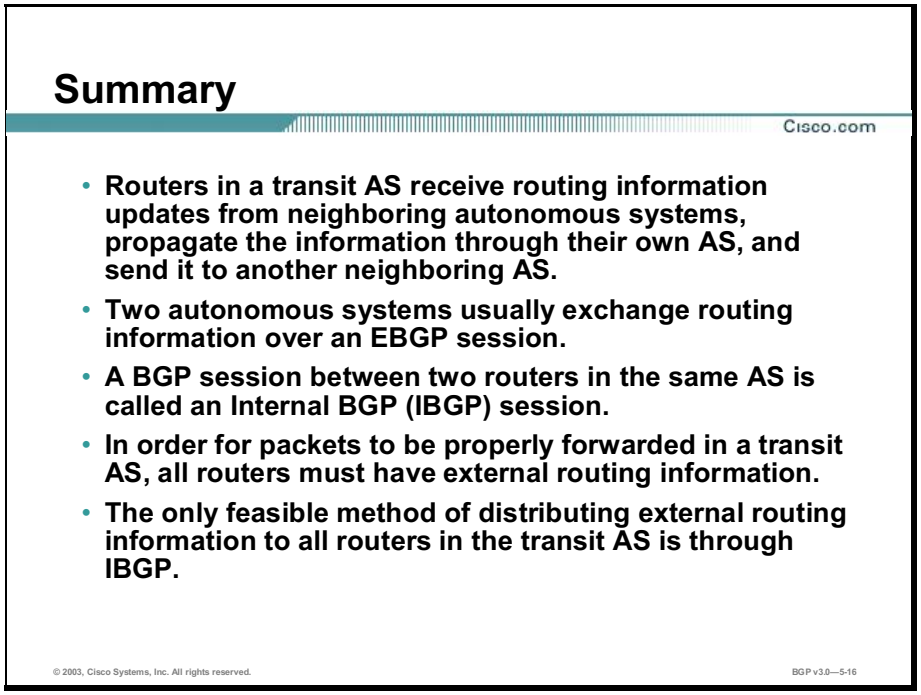
C) IGPs are not capable of handling the number of routes currently present on the Internet.

D) The increased convergence times of IGPs as compared to BGP would cause too many flaps, rendering BGP unstable for Internet use.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Routers in a transit AS receive routing information updates from neighboring autonomous systems, propagate the information through their own AS, and send it to another neighboring AS.
• Two autonomous systems usually exchange routing information over an EBGP session.
• A BGP session between two routers in the same AS is called an Internal BGP (IBGP) session.
• In order for packets to be properly forwarded in a transit AS, all routers must have external routing information.
• The only feasible method of distributing external routing information to all routers in the transit AS is through IBGP.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-16

Next Steps

After completing this lesson, go to:

- IBGP and EBGP Interaction in a Transit Autonomous System lesson

References

For additional information, refer to these resources:

- For more information on transit autonomous systems, refer to “BGP Case Studies Section 1” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html
- For more information on the attributes in transit autonomous systems, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

Quiz: Transit Autonomous System Functions

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- List the functions of a transit AS
- Describe external route propagation between autonomous systems in a BGP network
- Describe internal route propagation within a BGP AS
- Explain how transiting packets are forwarded inside a transit AS
- Explain the need for deploying IBGP on all core routers

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) Why is IBGP a mandatory component of a transit AS?

A) It is the only feasible way to ensure that all routers in the AS have consistent external routing information.

B) It eliminates the scalability issues of running an IGP within the transit AS.

C) Running IBGP on all routers is the only way to satisfy the filtering requirements of the transit AS.

D) An IGP is not capable of handling the potential routing loops in the transit AS.

 Copyright © 2003, Cisco Systems, Inc.

Q2) How is EBGP used in a transit AS?

A) as a means of transporting customer routes across the transit backbone

B) to exchange routes between different autonomous systems and the transit AS

C) to enhance scalability by transporting IGP routes for the transit AS

D) as a means of injecting local routes into the transit backbone

Q3) Why is it not recommended to redistribute BGP routes into an IGP for use in a transit backbone?

A) Redistribution removes all BGP attributes needed to ensure optimal routing within the transit AS.

B) An IGP cannot enforce complex administrative policies and route selection rules.

C) IGPs cannot scale to the demands presented by the number of routes on the Internet.

D) IGPs are not stable when faced with a flapping network.

Q4) What are the two key functions of a transit AS? (Choose two.)

A) to filter out routes that do not belong to customers of the service provider

B) to provide Internet connectivity to customers of the service provider

C) to propagate routes between remote autonomous systems

D) to route packets between remote networks

Q5) How are BGP routes sent across the transit backbone?

A) by redistributing BGP into an IGP and then back into BGP

B) through the use of IBGP

C) by establishing EBGP sessions between all routers in the transit backbone

D) by redistributing connected routes at the edge of the transit backbone

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

IBGP and EBGP Interaction in a Transit Autonomous System

Overview

A transit autonomous system (AS), by definition, allows traffic not originating in that AS to travel through it. This activity requires interaction between External Border Gateway Protocol (EBGP) and Internal Border Gateway Protocol (IBGP) in the transit AS. This lesson introduces the requirements of IBGP and how routers residing in the transit AS process the next-hop attribute. Changes to the normal processing of the next-hop attribute are also described in this lesson. The lesson concludes with a comparison between EBGP and IBGP.

Importance

Configuring a Border Gateway Protocol (BGP) network in a transit services configuration requires special care to ensure consistency of routing information throughout the AS. Understanding the interaction between EBGP and IBGP is crucial to successfully configuring and troubleshooting the transit autonomous network.

Objectives

Upon completing this lesson, you will be able to:

- Describe AS-path processing in IBGP
- Explain the need for BGP split horizon
- Explain the need for a full-mesh topology between IBGP routers, and its implications
- Explain the benefits of establishing IBGP neighbor sessions using loopback interfaces
- Describe next-hop processing in IBGP

- Explain why all EBGP peers must be reachable by all BGP-speaking routers within the AS
- Describe how to configure routers to announce themselves as the next hop in IBGP updates
- Explain the differences between EBGP and IBGP sessions

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

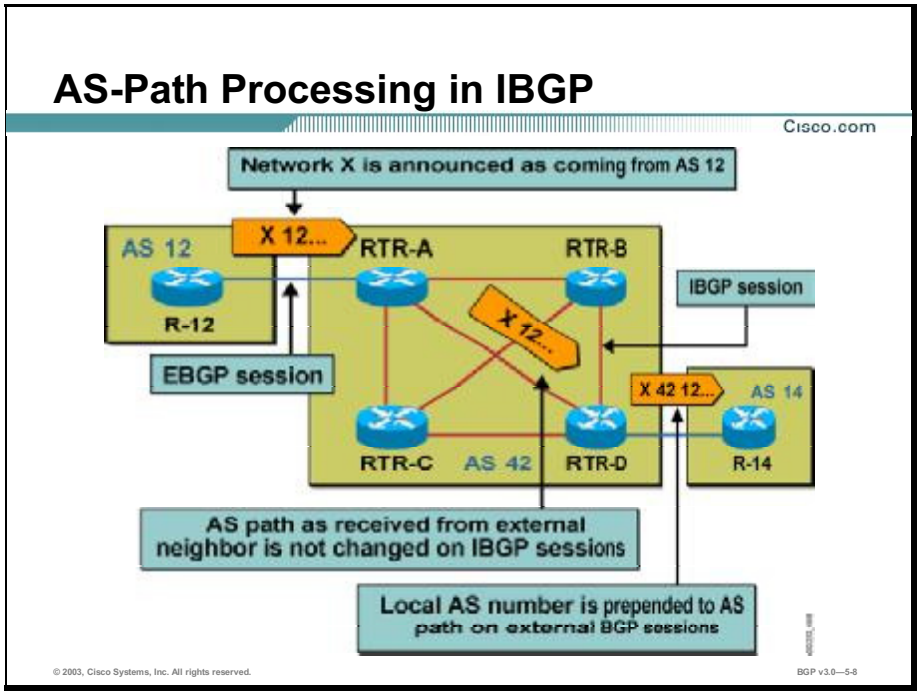
This lesson includes these topics:

- Overview
- AS-Path Processing in IBGP
- BGP Split Horizon
- IBGP Full Mesh
- IBGP Neighbors
- IBGP Next-Hop Processing
- Transit Network Using External Next Hops
- Transit Network Using Edge Routers as Next Hops
- Differences Between EBGP and IBGP Sessions
- Summary
- Assessment (Quiz): IBGP and EBGP Interaction in a Transit Autonomous System

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Processing in IBGP

This topic describes AS path processing in IBGP.



All BGP routing updates carry the mandatory well-known attribute AS-path, which lists the autonomous systems that the routing update has already crossed.

When a router originates a BGP prefix (network X in this example), the AS path is empty. Whenever a BGP prefix is announced over an EBGP session, the AS number of the router sending the information is prepended to the AS path. In the example, R-12 inserts “12” in the AS path before forwarding the routing update to RTR-A.

The AS path is not changed when the BGP prefix is propagated across IBGP sessions because the routing update has not crossed an AS boundary. In the example, RTR-A forwards the information over an IBGP session to RTR-D with the AS path unchanged. The AS-path information about network X will be the same in all routers within AS 42, because all the routers are updated using IBGP sessions from RTR-A.

When RTR-D forwards the information about network X to R-14, it prepends its own AS number (42) to the AS path. Thus, R-14 receives the routing information about network X with an AS-path attribute of “42 12.”

Practice

Q1) How is the AS-path attribute modified across a transit AS?

A) The AS path of each router is prepended to the AS-path attribute.

B) The AS path is not changed when the BGP prefix is propagated across IBGP sessions.

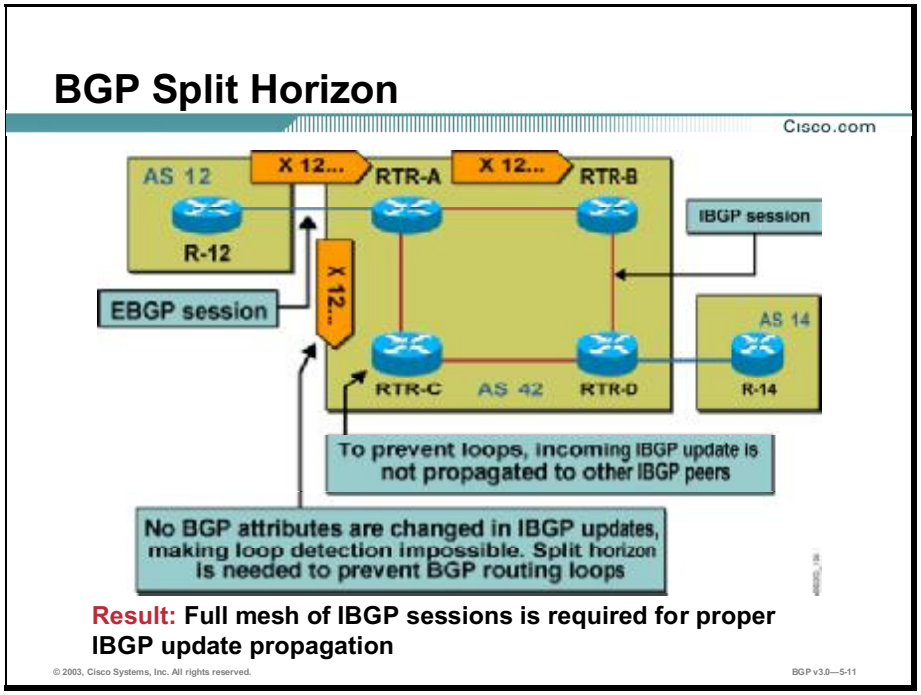
C) The AS number of the router sending the information is appended to the AS path.

D) The value of the AS path is reset to the AS path of the transit backbone.

 Copyright © 2003, Cisco Systems, Inc.

BGP Split Horizon

This topic explains the need for BGP split horizon as a mechanism to prevent routing loops.


BGP Split Horizon
Cisco.com
AS 12
X 12...
RTR-A
X 12...
RTR-B
IBGP session
R-12
EBGP session
X 12...
AS 14
RTR-C
AS 42
RTR-D
R-14
To prevent loops, incoming IBGP update is not propagated to other IBGP peers
No BGP attributes are changed in IBGP updates, making loop detection impossible. Split horizon is needed to prevent BGP routing loops
Result: Full mesh of IBGP sessions is required for proper IBGP update propagation
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-11

All routers within an AS must make routing decisions in a consistent way. They must have access to the same routing information with the same attributes in order to come to the same conclusion about which exit point of the AS to use. In other words, the BGP attributes should not be changed within the AS.

The AS-path attribute is not changed over an IBGP session, because the BGP update has not crossed the AS boundary. However, the AS-path attribute is the primary means of detecting routing information loops—a BGP router that encounters its own AS in the AS path of an incoming BGP update silently ignores the information. Because the AS path is modified by BGP speaking routers only on EBGP sessions, this loop-preventing mechanism is useful between autonomous systems only, not within them.

Routing information loops within the AS are prevented by IBGP split horizon—routing information received through an IBGP session is never forwarded to another IBGP neighbor, only toward EBGP neighbors. Due to BGP split horizon, no router can relay IBGP information within the AS—all routers must be directly updated from the border router that received the EBGP update.

Practice

Q1) What is the purpose of the BGP split-horizon rule?

A) to propagate IBGP updates to other IBGP peers

B) to prevent BGP routing loops within an AS

C) to send routing information to an EBGP neighbor

D) to prevent BGP routes from being sent back to the advertising router

Q2) What is the impact of BGP split horizon on an AS?

A) IBGP-speaking routers forward BGP updates to all other IBGP-speaking routers within the transit backbone.

B) The IBGP full-mesh requirement within the transit AS is no longer required because BGP routing loops are eliminated.

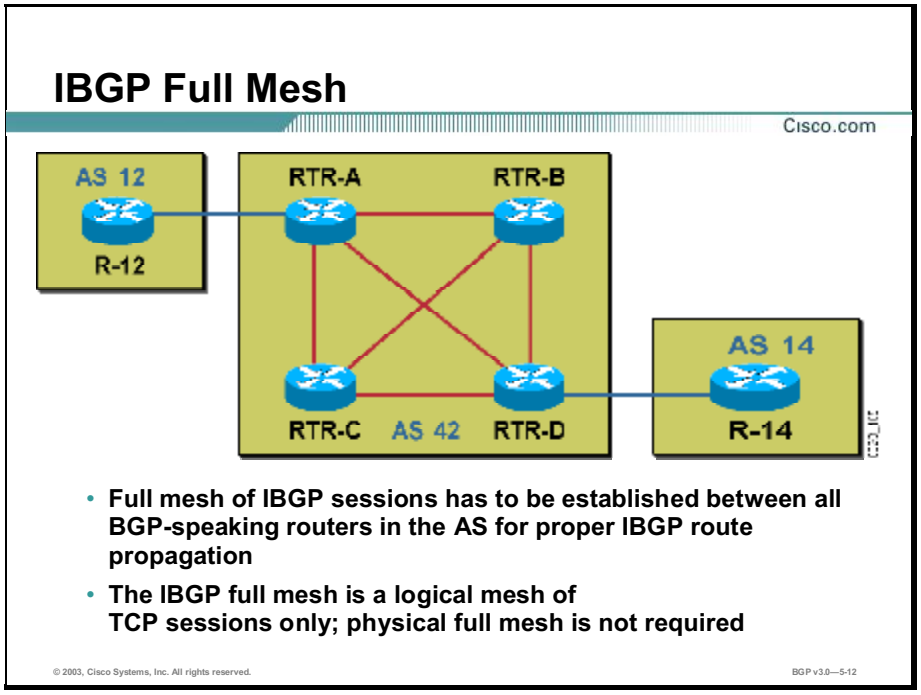
C) Convergence time within the AS is reduced because fewer routes have to be processed by each IBGP neighbor router.

D) All BGP routers in an AS must be updated directly by the border router receiving the update via EBGP.

 Copyright © 2003, Cisco Systems, Inc.

IBGP Full Mesh

This topic explains the need for a full-mesh topology between IBGP routers, and its implications.


IBGP Full Mesh
Cisco.com
AS 12
R-12
RTR-A
RTR-B
RTR-C
AS 42
RTR-D
AS 14
R-14
• Full mesh of IBGP sessions has to be established between all BGP-speaking routers in the AS for proper IBGP route propagation
• The IBGP full mesh is a logical mesh of TCP sessions only; physical full mesh is not required
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-12

Because every router on the transit path within the AS must have routing information about all external networks received by any of the border routers, all routers must have IBGP sessions to all border routers. But this is not enough, because any of the internal routers could also create new BGP routing information (for example, originate a customer network). These updates must also reach all the routers within the AS. The conclusion is that all BGP routers within an AS must have IBGP sessions with every other BGP router in the AS, resulting in a full mesh of BGP sessions between BGP-speaking routers in an AS.

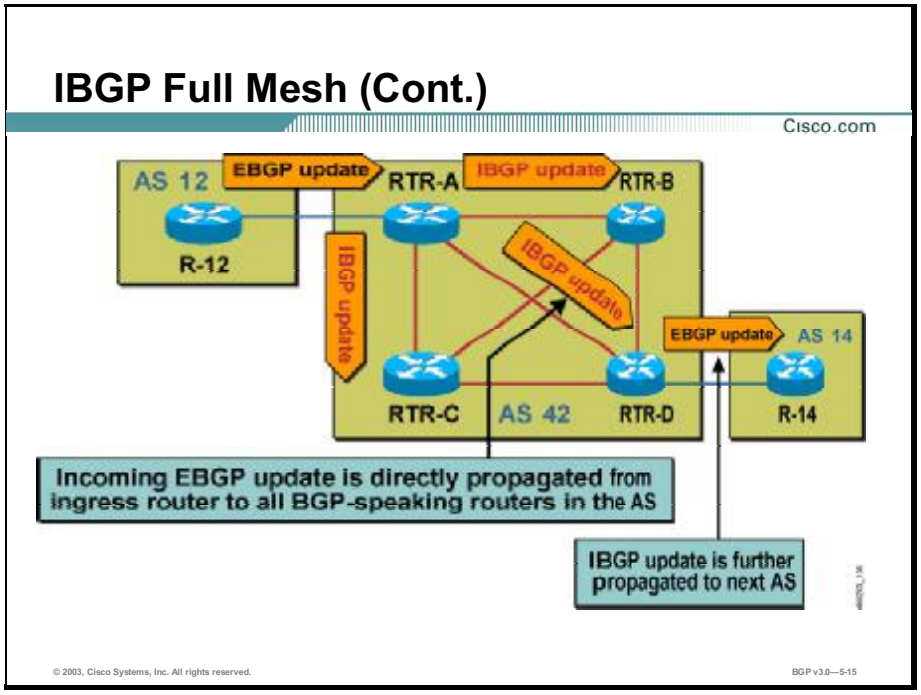
In the sample network, RTR-A must have IBGP sessions with RTR-B, RTR-C, and RTR-D in order to propagate routes received from AS 12 to all routers within AS 42. Similarly, RTR-D must have IBGP sessions with RTR-A, RTR-B, and RTR-C to be able to propagate routes received from AS 14 to all routers within AS 42.

Note	The IBGP session between RTR-B and RTR-C is not strictly necessary for proper forwarding of IP packets between external destinations. It does become mandatory if RTR-B or RTR-C starts to originate BGP networks. To prevent potential future connectivity issues, it is a good practice to establish a full mesh of IBGP sessions regardless of whether they are needed at the time of network deployment or not.

The IGP that runs within AS 42 provides enough information to any BGP router within AS 42 to send IP packets to any other router in the AS. Having enough router reachability information makes it possible to establish IBGP sessions between routers even though they are not physically connected. The IBGP full mesh is a logical full mesh of TCP sessions and will run on an arbitrary physical topology.

 Copyright © 2003, Cisco Systems, Inc.

Example



The figure here illustrates IBGP split-horizon and IBGP full-mesh principles in the sample network. R-12 is sending an update to RTR-A over an EBGP session. Updates received on an EBGP session should be forwarded on all other IBGP sessions, so RTR-A updates RTR-B, RTR-C, and RTR-D. All routers within AS 42 are updated directly by RTR-A.

RTR-B and RTR-C are prevented from forwarding the update that they received from RTR-A due to BGP split horizon.

RTR-D, which received the information on an IBGP session, is prevented from updating RTR-B and RTR-C due to the same split-horizon rule. But RTR-D will update R-14 over an EBGP session.

Practice

Q1) What are two reasons why a fully meshed logical topology is required for all IBGP neighbors in an AS? (Choose two.)

A) to receive routing information about external networks from border routers

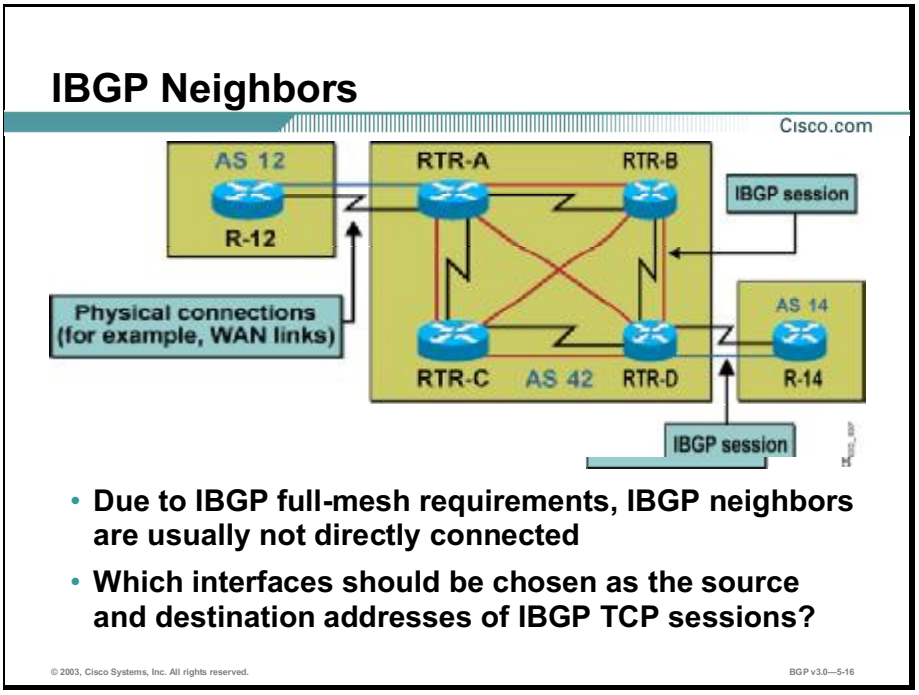
B) to facilitate BGP split-horizon rules

C) to allow IBGP-speaking routers to inject routes into BGP

D) to ensure that each IBGP router has the most current BGP table and attributes

IBGP Neighbors

This topic explains the benefits of establishing IBGP neighbor sessions using loopback interfaces.

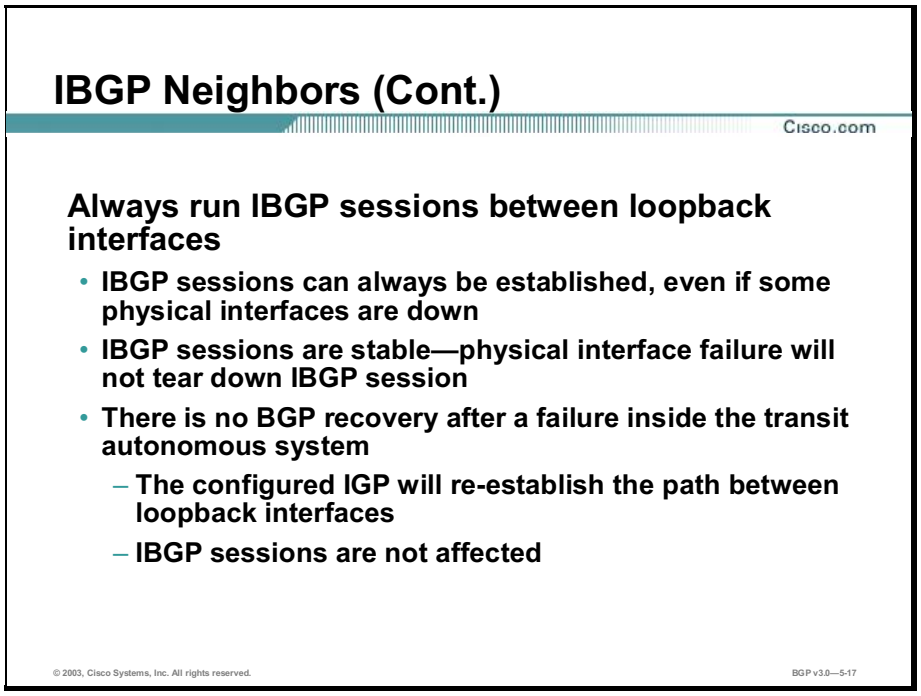


In this example, the transit AS 42 has a redundant physical topology. The IGP provides reachability information for all routers and networks within AS 42, allowing all routers in the AS to establish IBGP sessions to all other routers, even if not directly connected.

If the IBGP session between RTR-A and RTR-D was established using IP addresses that belong to the physical WAN interfaces, the IBGP session would go down if either of the WAN interfaces went down. As a result, the router would tear down the TCP session used for BGP between the routers because the IP address of an interface that is in the down state is invalid. Subsequently, all IP packets received with a destination address pointing to that interface will also be dropped.

Network designers must be careful during the network design and implementation phase that those IBGP sessions remain established as long as the two BGP routers have any usable path between them.

 Copyright © 2003, Cisco Systems, Inc.


IBGP Neighbors (Cont.)
Cisco.com
Always run IBGP sessions between loopback interfaces
• IBGP sessions can always be established, even if some physical interfaces are down
• IBGP sessions are stable—physical interface failure will not tear down IBGP session
• There is no BGP recovery after a failure inside the transit autonomous system
– The configured IGP will re-establish the path between loopback interfaces
– IBGP sessions are not affected
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-17

The best choice when you are configuring IBGP sessions is to establish each session between loopback interfaces on each BGP router.

In order to establish BGP connectivity between the loopback interfaces, the IP addresses of these interfaces have to be reachable by both routers. It is important that the IGP carry information about the subnets assigned to each loopback interface so that they are reachable by all BGP routers in the AS.

The IBGP sessions established between loopback interfaces have increased stability. These sessions will not go down if a single physical interface goes down. As long as the IGP can find any path between the two routers, the IBGP session will remain up. BGP will not notice that the IGP has changed the traffic path between the two routers.

Note	Because BGP sessions run over TCP, BGP sessions can survive even a short loss of connectivity between BGP routers with no impact to the BGP routing protocol. The only requirement placed on the IGP is that the network must converge before the BGP keepalive timer expires.

Practice

Q1) What is the recommended way to run IBGP sessions?

A) Establish IBGP sessions using IP addresses of physical interfaces to allow detection of link failures.

B) Establish an IBGP session using IP addresses of loopback interfaces to prevent IBGP sessions from failing if a physical interface fails.

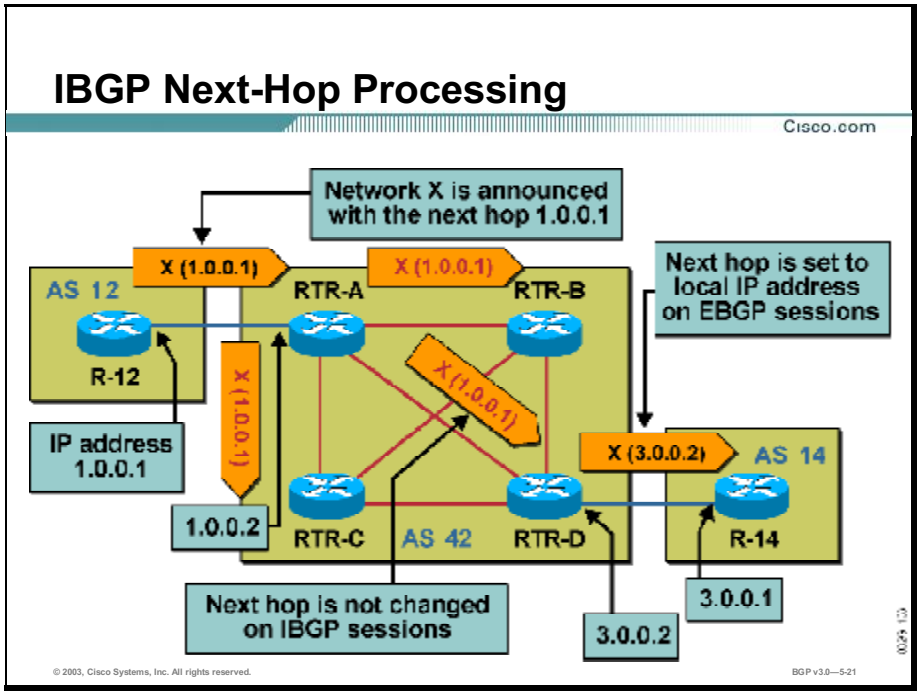
C) Establish IBGP sessions using the IP addresses of the fastest interface to make sure that packets are forwarded out the fastest interface.

D) There is no recommendation. Any IP address on a router can be used for an IBGP session without any penalties.

 Copyright © 2003, Cisco Systems, Inc.

IBGP Next-Hop Processing

This topic describes how the next-hop attribute is processed in internal BGP.



Every BGP update carries the mandatory well-known attribute next-hop, which specifies the IP address that should be used by the router as the forwarding next hop for packets sent toward the announced destination address. In most cases, the next hop is set to the IP address that the sending router is using as its source IP address for EBGP sessions. The receiving BGP router will use the information and route IP packets toward the announced destination via the indicated next hop, which is normally directly connected.

The next-hop attribute is not changed on IBGP updates, meaning that when the border router forwards the BGP update on IBGP sessions, the next-hop address is still set to the IP address of the far end of the EBGP session. Therefore, the receiver of IBGP updates will see the next-hop information indicating a destination that is not directly connected. To resolve this problem, the router will check its forwarding table and see if and how it can reach the next-hop address. The router can then route IP packets with destination addresses matching the network in the BGP update in the same direction as it would have routed an IP packet with a destination address equal to the IP address stated in the next-hop attribute. This process is known as recursive routing.

In the figure, R-12 sends a BGP update about network X. Because it is sending this update over an EBGP session to RTR-A, the next-hop attribute is set to the IP address used at the R-12 side of the EBGP session, 1.0.0.1.

RTR-A can use this information and route packets to network X by forwarding them to R-12.

RTR-A also forwards the BGP update over all its IBGP sessions. It does not change the next-hop attribute, so RTR-B, RTR-C, and RTR-D will get information that they can reach network

X by forwarding packets to 1.0.0.1. But that IP address is not directly connected, so the routers must look in their forwarding tables to see if and how they can reach 1.0.0.1. If the recursive route lookup is successful, each router can then route packets to network X in the same direction as they would route packets to 1.0.0.1.

RTR-D also forwards the BGP update about network X to R-14. The connection between these routers is an EBGP session, meaning that RTR-D will set the next-hop attribute to its own IP address, 3.0.0.2, which is used by RTR-D on the EBGP session toward R-14.

Practice

Q1) How is the BGP next hop changed inside an AS?

A) The next-hop address is set to the router-ID of the forwarding router.

B) The next-hop address is not modified.

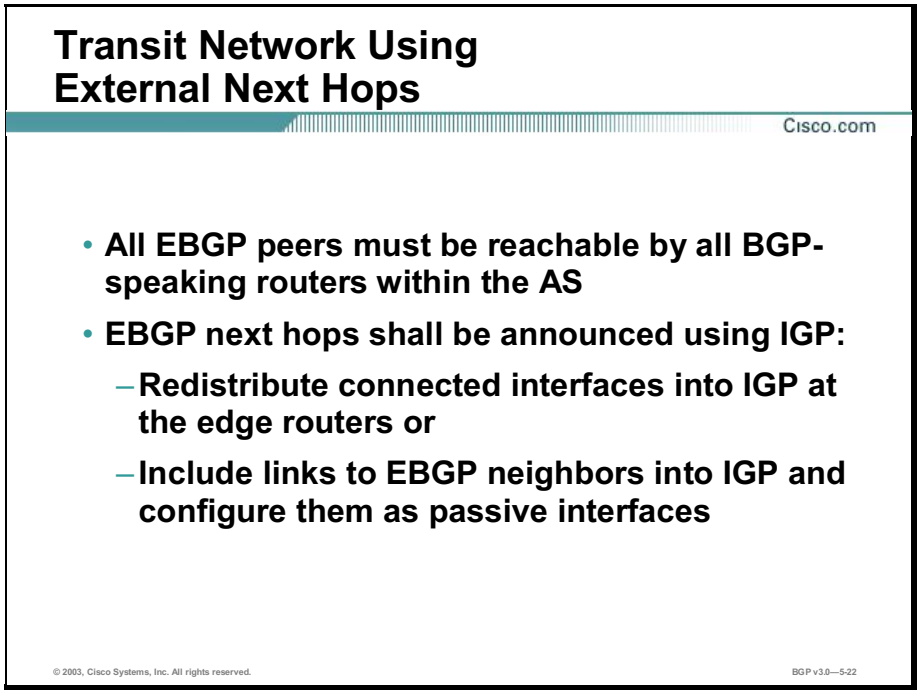
C) The next-hop address is set to the address of the receiver.

D) The next-hop attribute is set to the configured default gateway.

 Copyright © 2003, Cisco Systems, Inc.

Transit Network Using External Next Hops

This topic explains why all EBGP peers must be reachable by all BGP-speaking routers within the transit AS.


Transit Network Using External Next Hops
Cisco.com
• All EBGP peers must be reachable by all BGP-speaking routers within the AS
• EBGP next hops shall be announced using IGP:
– Redistribute connected interfaces into IGP at the edge routers or
– Include links to EBGP neighbors into IGP and configure them as passive interfaces
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-22

All BGP-speaking routers within the AS get information about external networks with the next-hop attribute, which is set to the far end of the EBGP sessions reaching the border routers of the AS.

Routers use a recursive routing mechanism when they determine how to forward IP packets toward external destinations. When BGP routes are used in the forwarding table, the router checks how it would have reached the next-hop address, and it installs the BGP route with the same forwarding indication as for the route used to reach the next-hop IP address.

In order to get the recursive routing to work, the router must resolve all possible next-hop references using information in the forwarding table, which is already there. The IGP used within the AS must carry this information.

One way of making the IGP carry the information necessary to resolve the BGP next-hop addresses is to make sure that all the border routers, which contain the EBGP sessions, redistribute connected subnets into the IGP using the **redistribute connected** routing protocol configuration command. Because EBGP sessions are established between routers using a directly connected interface, the far end of the EBGP sessions is an IP address within the directly connected subnet. By redistributing the connected interfaces into the IGP, the border routers allow next-hop references to be resolvable by all routers within the AS.

External subnets redistributed into the IGP might appear as external IGP routes, depending on what IGP is configured within the AS. There exist several scalability issues associated with external routes in some routing protocols (for example, Open Shortest Path First [OSPF] carries

each external subnet in a separate link-state advertisement [LSA] object). If route redistribution is not desirable for any reason, an alternative method is to include the subnet on which the EBGP session is running in the IGP configuration using the **network** command. To prevent the border router from exchanging IGP routing with the border router of the other AS, you must configure the interface as a passive interface. Failure to do so could cause the two different autonomous systems to exchange routes using the IGP. In that case, all benefits of having separate autonomous systems will be lost.

Practice

Q1) What are the implications of IBGP next-hop processing on the network design?

A) The network must contain a physical full mesh between all BGP-speaking routers.

B) All BGP external routes must be redistributed into the IGP by the border router.

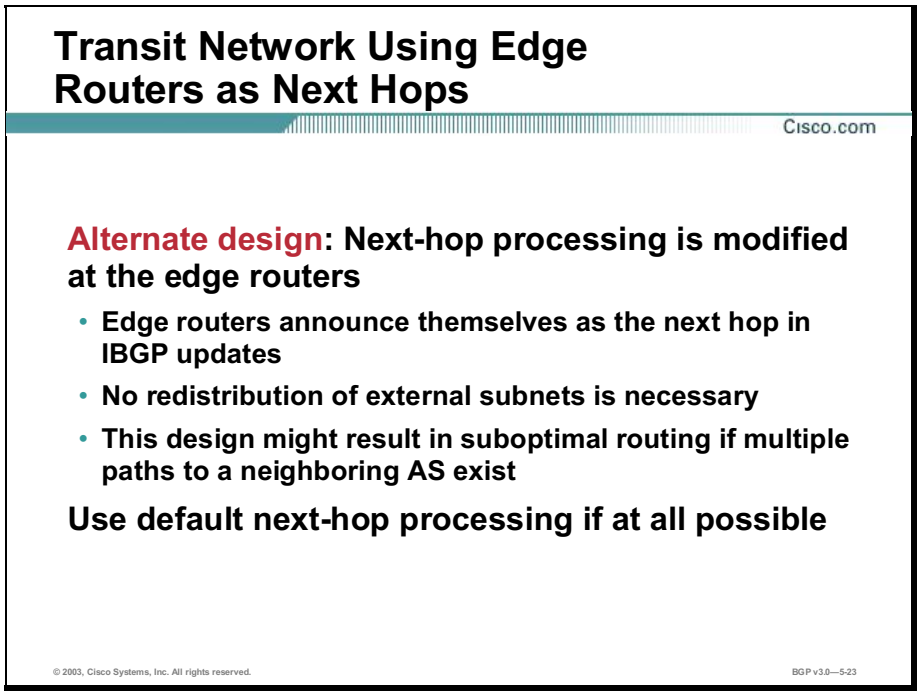
C) All BGP speakers must have IGP reachability to all external neighbors to be able to perform a recursive lookup to resolve next-hop addresses.

D) You must configure static routes on all IBGP-speaking routers pointing to the border routers of the transit AS.

 Copyright © 2003, Cisco Systems, Inc.

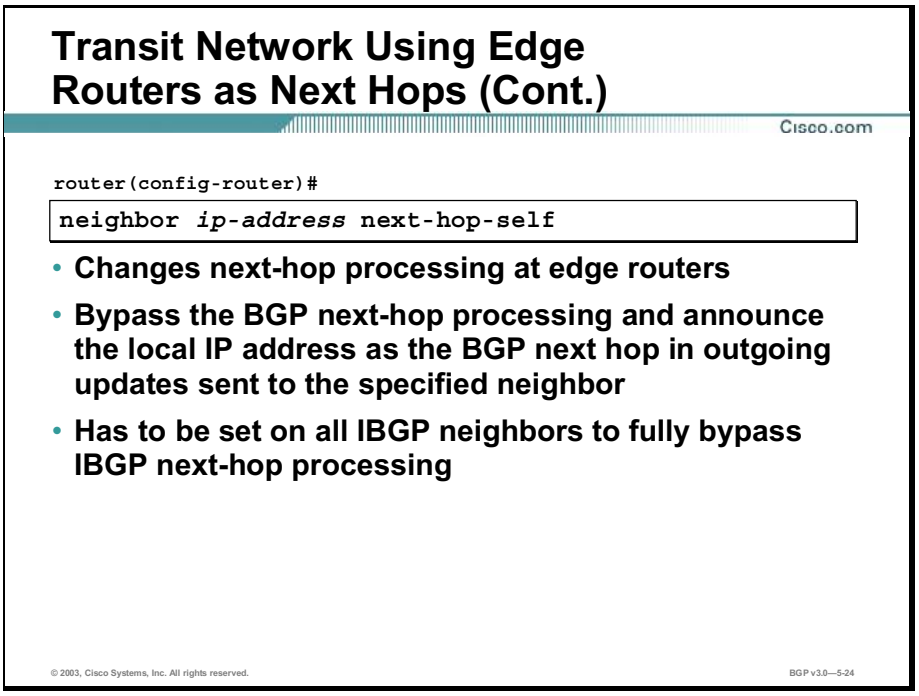
Transit Network Using Edge Routers as Next Hops

This topic describes how to configure edge routers to announce themselves as the next hop in IBGP updates.


Transit Network Using Edge Routers as Next Hops
Cisco.com
Alternate design: Next-hop processing is modified at the edge routers
• Edge routers announce themselves as the next hop in IBGP updates
• No redistribution of external subnets is necessary
• This design might result in suboptimal routing if multiple paths to a neighboring AS exist
Use default next-hop processing if at all possible
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-23

The next-hop attribute is usually not modified by an IBGP peer when the BGP update is propagated across IBGP sessions. However, you could configure the BGP router to have a different behavior and set its IP address as the next-hop address even when the BGP updates are sent across IBGP sessions (emulating behavior on EBGP sessions). If you do configure an IBGP router to emulate the behavior of EBGP sessions on the IBGP sessions of the border routers, the BGP updates received on the EBGP sessions will be forwarded on the IBGP sessions and the next-hop attribute will be set to the IP address used on the local side of the IBGP session. The original next hop, set by the far end of the EBGP session, will be lost.

The receiver of the IBGP information will do recursive routing the normal way. But the next-hop address used will be the IP address of the far end of the IBGP session, because the border router changed it. The IP address of the far-end IBGP peer is always known in the forwarding table; otherwise, the IBGP session would not have been established. There is no need for the receiver of the IBGP information to have knowledge of how to reach the far end of the EBGP session, because that IP address is no longer set as the next hop.


Transit Network Using Edge Routers as Next Hops (Cont.)
Cisco.com
router(config-router)#
neighbor ip-address next-hop-self
• Changes next-hop processing at edge routers
• Bypass the BGP next-hop processing and announce the local IP address as the BGP next hop in outgoing updates sent to the specified neighbor
• Has to be set on all IBGP neighbors to fully bypass IBGP next-hop processing
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-24

neighbor next-hop-self

To configure the router as the next hop for a BGP-speaking neighbor or peer group, use the **neighbor next-hop-self** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **next-hop-self**

To disable this feature, use the **no** form of this command.

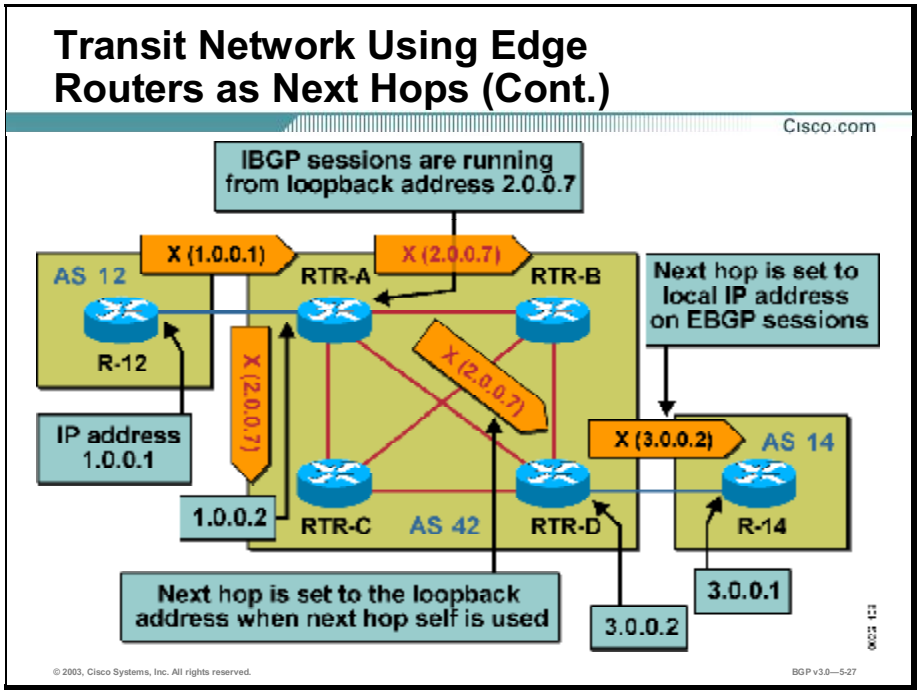
no neighbor {*ip-address* | *peer-group-name*} **next-hop-self**

Syntax Description

ip-address	IP address of the BGP-speaking neighbor
peer-group-name	Name of a BGP peer group

 Copyright © 2003, Cisco Systems, Inc.

Example



The next-hop attribute is normally not changed on IBGP updates. But in the figure, the **next-hop-self** configuration has been used on all IBGP sessions. When the border router forwards the incoming EBGP update over an outgoing IBGP session, the border router changes the next-hop address to the IP address used as the source address of the IBGP session.

The receiver of IBGP updates will see next-hop information indicating a destination, which might not be directly connected. To resolve this problem, it will check its forwarding table and see if and how the next-hop address can be reached. Then it will route IP packets with destination addresses matching the network in the BGP update in the same direction as it would have routed an IP packet with the destination address equal to the IP address stated in the next-hop attribute. In this case, it is obvious that the next-hop address can be reached, because the IBGP session would not be established otherwise.

In the figure, R-12 sends a BGP update about network X. Because it is sending a BGP update over an EBGP session to RTR-A, the next-hop attribute is set to the IP address used at the R-12 side of the EBGP session, 1.0.0.1.

RTR-A can use this information and route packets to network X by forwarding them to R-12.

RTR-A also forwards the BGP update on all its IBGP sessions. It changes the next-hop attribute to the IP address of its own loopback interface, so RTR-B, RTR-C, and RTR-D will get information that they can reach network X by forwarding packets to 2.0.0.7. But that address is not directly connected. The routers will inspect the forwarding table to see if and how they can reach 2.0.0.7. They can then route packets to network X in the same direction as they would route packets to 2.0.0.7.

RTR-D also forwards the BGP update about network X to R-14. This is an EBGP session, which means that RTR-D will set the next-hop attribute to its own IP address used on that EBGP session, 3.0.0.2.

Practice

Q1) How can you influence IBGP next-hop processing?

A) by configuring the router to set the next-hop attribute to its router-ID

B) by configuring the router to set the next-hop attribute to the gateway of last resort

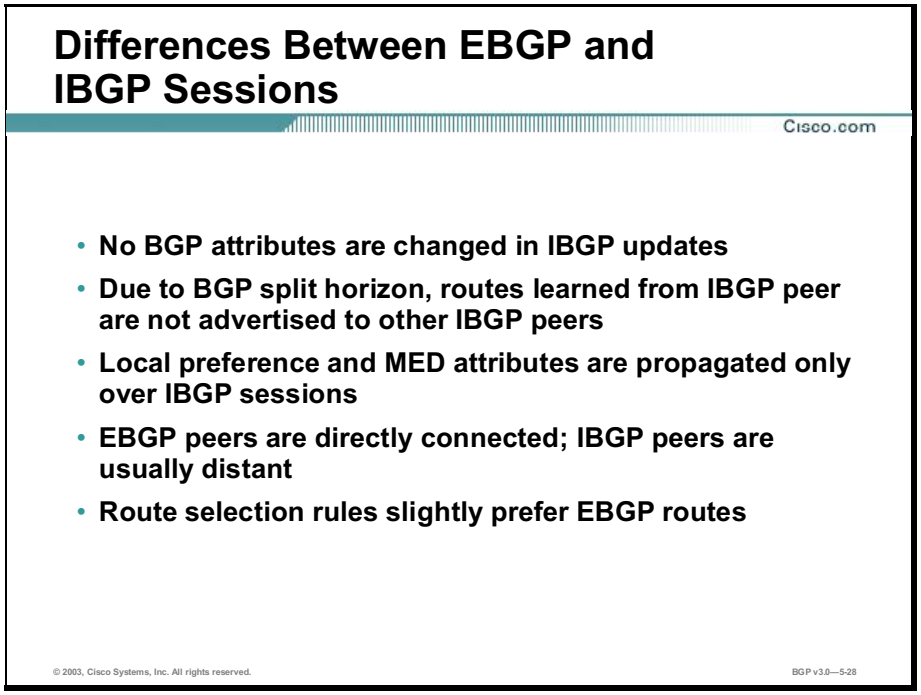
C) by configuring the router to ignore the next-hop attribute

D) by configuring the router to set the next-hop attribute to its source address of internal BGP sessions

 Copyright © 2003, Cisco Systems, Inc.

Differences Between EBGP and IBGP Sessions

This topic explains the differences between EBGP and IBGP sessions.


Differences Between EBGP and IBGP Sessions
Cisco.com
• No BGP attributes are changed in IBGP updates
• Due to BGP split horizon, routes learned from IBGP peer are not advertised to other IBGP peers
• Local preference and MED attributes are propagated only over IBGP sessions
• EBGP peers are directly connected; IBGP peers are usually distant
• Route selection rules slightly prefer EBGP routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-28

Both EBGP and IBGP sessions forward BGP updates. But they do it in a slightly different manner:

- The router does not change BGP attributes when an update is sent across an IBGP session, unless **next-hop-self** is configured. When a BGP speaking router sends an update across an EBGP session, the next-hop attribute is always set and the AS number of the router is prepended to the AS-path attribute.

- IBGP uses split horizon to prevent routing information loops. EBGP does not use split horizon and instead uses the AS path to detect loops. In both cases, a router forwards only the best route and never sends a route back on the session from which it was received. But IBGP split-horizon rules also prohibit a router from forwarding any information received on an IBGP session to another IBGP session.

- IBGP border routers will remove the local preference attribute from a BGP route before the BGP update is sent over an EBGP session. This difference means that the local preference attribute is distributed on IBGP sessions only.

- The multi-exit discriminator (MED) attribute is received from a neighboring AS and distributed within the local AS. But when the update is about to be forwarded over another EBGP session to a third AS, the information is stripped off.

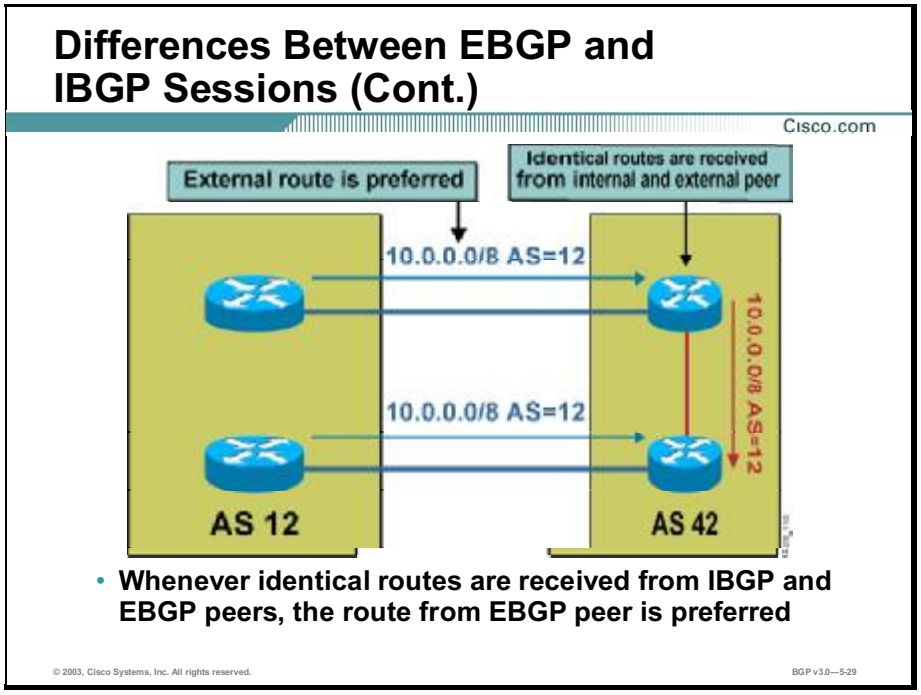
- Two routers with an EBGP session between them normally establish the session using the IP addresses from a common, shared subnet. Using the shared subnet to establish the session guarantees that the two routers can exchange IP packets without any IGP running between them. Also, recursive routing will always succeed because the next-hop address is reachable using a directly connected route.

- IBGP sessions are normally established between all routers in the AS in a full mesh. But all routers in an AS might not have physical connections to every other router within the AS. Since IBGP sessions are established between routers using IP addresses of different subnets, an IGP must be running within the AS in order to establish IBGP sessions.

- BGP route selection rules slightly favor EBGP routes over equivalent IBGP routes.

 Copyright © 2003, Cisco Systems, Inc.

Example


Differences Between EBGP and IBGP Sessions (Cont.)
Cisco.com
External route is preferred
Identical routes are received from internal and external peer
10.0.0.0/8 AS=12
10.0.0.0/8 AS=12
10.0.0.0/8 AS=12
AS 12
AS 42
• Whenever identical routes are received from IBGP and EBGP peers, the route from EBGP peer is preferred
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-29

One of the default goals of transit packet forwarding is to propagate the transit packet toward the downstream AS as soon as possible. A border router that receives otherwise equivalent routes to the same destination over both an EBGP session and an IBGP session will prefer the information received through the EBGP session.

Note	Equivalent routes are routes that have equal BGP path attributes used in BGP route selection rules (weight, local preference, AS-path length, origin, MED).

In the figure, the lower router in AS 42 receives BGP updates about network 10.0.0.0/8 over two different paths. One update is received over the EBGP session to AS 12. The other update is received over the IBGP session to the upper router in AS 42. All essential attributes are the same, so route selection cannot be made easily.

The lower router in AS 42 realizes that IP packets with destination addresses within network 10.0.0.0/8 should sooner rather than later leave AS 42. It is better to make them leave the AS right away. So the update received on the EBGP session is preferred over the update received on the IBGP session.

Practice

Q1) What are the three major differences that exist between EBGP and IBGP? (Choose three.)

A) No BGP attributes are changed in IBGP updates.

B) Route selection rules slightly prefer IBGP routes.

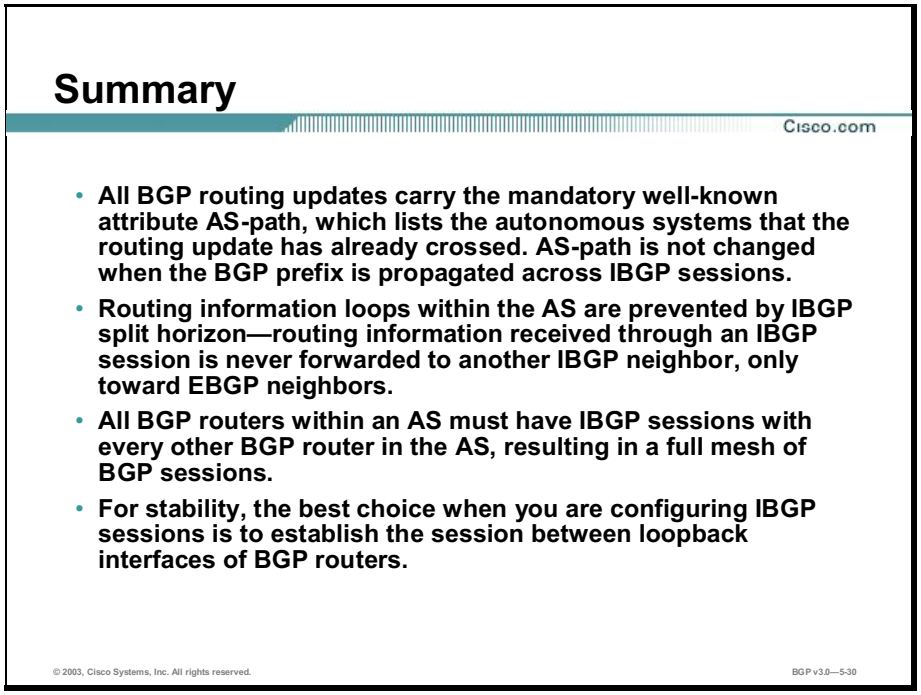
C) Local preference and MED are propagated only over IBGP sessions.

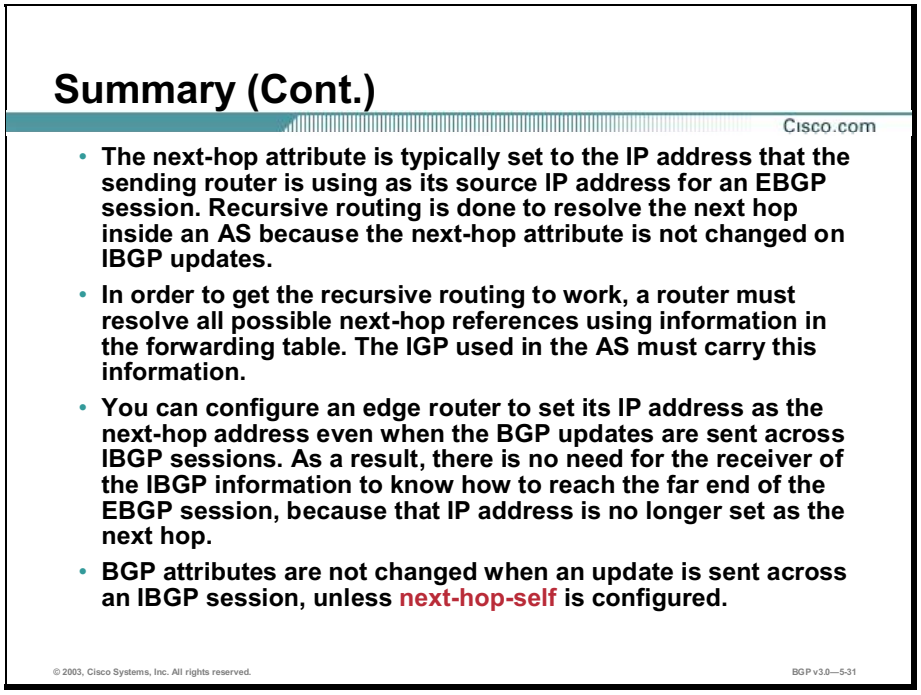
D) Routes learned from IBGP are not advertised to other IBGP peers.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• All BGP routing updates carry the mandatory well-known attribute AS-path, which lists the autonomous systems that the routing update has already crossed. AS-path is not changed when the BGP prefix is propagated across IBGP sessions.
• Routing information loops within the AS are prevented by IBGP split horizon—routing information received through an IBGP session is never forwarded to another IBGP neighbor, only toward EBGP neighbors.
• All BGP routers within an AS must have IBGP sessions with every other BGP router in the AS, resulting in a full mesh of BGP sessions.
• For stability, the best choice when you are configuring IBGP sessions is to establish the session between loopback interfaces of BGP routers.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-30


Summary (Cont.)
Cisco.com
• The next-hop attribute is typically set to the IP address that the sending router is using as its source IP address for an EBGP session. Recursive routing is done to resolve the next hop inside an AS because the next-hop attribute is not changed on IBGP updates.
• In order to get the recursive routing to work, a router must resolve all possible next-hop references using information in the forwarding table. The IGP used in the AS must carry this information.
• You can configure an edge router to set its IP address as the next-hop address even when the BGP updates are sent across IBGP sessions. As a result, there is no need for the receiver of the IBGP information to know how to reach the far end of the EBGP session, because that IP address is no longer set as the next hop.
• BGP attributes are not changed when an update is sent across an IBGP session, unless next-hop-self is configured.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-31

Next Steps

After completing this lesson, go to:

- Packet Forwarding in Transit Autonomous Systems lesson

References

For additional information, refer to these resources:

- For more information on transit autonomous systems, refer to “BGP Case Studies Section 1” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html

- For more information on the next-hop attribute, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

 Copyright © 2003, Cisco Systems, Inc.

Quiz: IBGP and EBGP Interaction in a Transit Autonomous System

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe AS-path processing in IBGP
- Explain the need for BGP split horizon
- Explain the need for a full-mesh topology between IBGP routers, and its implications
- Explain the benefits of establishing IBGP neighbor sessions using loopback interfaces
- Describe next-hop processing in IBGP
- Explain why all EBGP peers must be reachable by all BGP-speaking routers within the AS
- Describe how to configure routers to announce themselves as the next hop in IBGP updates
- Explain the differences between EBGP and IBGP sessions

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) Which two statements are true regarding the AS-path attribute as it relates to IBGP? (Choose two.)

A) Each router in the AS appends its AS number to the AS path on outgoing BGP updates.

B) The AS path inside an AS will be empty for routes originating inside a neighboring AS.

C) The AS-path attribute is not used to detect routing loops inside an AS.

D) The AS-path attribute is not modified within the AS.

Q2) Why is it recommended that loopback interfaces be used to form IBGP neighbor sessions?

A) Using loopback interfaces reduces router memory resource requirements.

B) Using loopback interfaces reduces router CPU resource requirements.

C) Using loopback interfaces ensures IBGP session stability.

D) Using loopback interfaces is more secure than using the physical interface.

Q3) How is the BGP next-hop attribute processed over an IBGP connection?

A) The next-hop address is set to the address of the receiving router.

B) The next-hop address is not modified over the IBGP session.

C) The next-hop address is set to the IP address of the nearest EBGP peer.

D) The next-hop attribute is set to the IP address of the nearest EBGP peer; if no external AS connection has been configured, the next hop is set to the default gateway configured on the router.

Q4) Which two statements are true of the full-mesh requirement in IBGP? (Choose two.)

A) The IBGP mesh requires a logical full mesh.

B) A physical full mesh must be maintained within the IBGP AS.

C) Due to BGP split horizon, no router can relay IBGP information within the AS.

D) All routers within the AS must be directly connected to ensure correct delivery of BGP routing information.

 Copyright © 2003, Cisco Systems, Inc.

Q5) What three statements are true regarding the next-hop-self configuration in BGP? (Choose three.)

A) Changing the next-hop attribute might cause suboptimal routing.

B) This configuration changes how the next-hop attribute is processed at edge routers.

C) This configuration announces the local IP address as the BGP next hop in outgoing updates sent to the specified neighbor.

D) This configuration removes the requirement for the IGP to carry reachability information for intra-AS destinations.

Q6) What are three differences between IBGP and EBGP sessions? (Choose three.)

A) Route selection rules slightly prefer IBGP routes.

B) Routes learned from IBGP peer are not advertised to other IBGP peers.

C) EBGP peers are directly connected, and IBGP peers are usually distant.

D) By default, no BGP attributes are changed in IBGP updates.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Packet Forwarding in Transit Autonomous Systems

Overview

A transit autonomous system (AS) requires interaction between External Border Gateway Protocol (EBGP) and Internal Border Gateway Protocol (IBGP), and between IBGP and an Interior Gateway Protocol (IGP) in the transit AS. This lesson describes packet forwarding through a transit AS and discusses the requirements for successful packet forwarding, such as recursive route lookup and an IGP in the transit AS. This lesson concludes with a discussion of the interaction between IBGP and an IGP running within the transit AS.

Importance

Configuring a Border Gateway Protocol (BGP) network in a transit services configuration requires special care to ensure the consistency of routing information throughout the AS. Understanding the interaction between IBGP and EBGP, and between IBGP and an IGP is crucial to successfully configuring and troubleshooting the transit autonomous network.

Objectives

Upon completing this lesson, you will be able to:

- Describe packet forwarding in a transit AS
- Explain recursive lookup in Cisco IOS® software
- Explain the need for an IGP in a transit backbone running BGP on all routers
- Describe interactions between BGP and IGP
- Explain the potential problems that might arise from BGP and IGP interaction

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

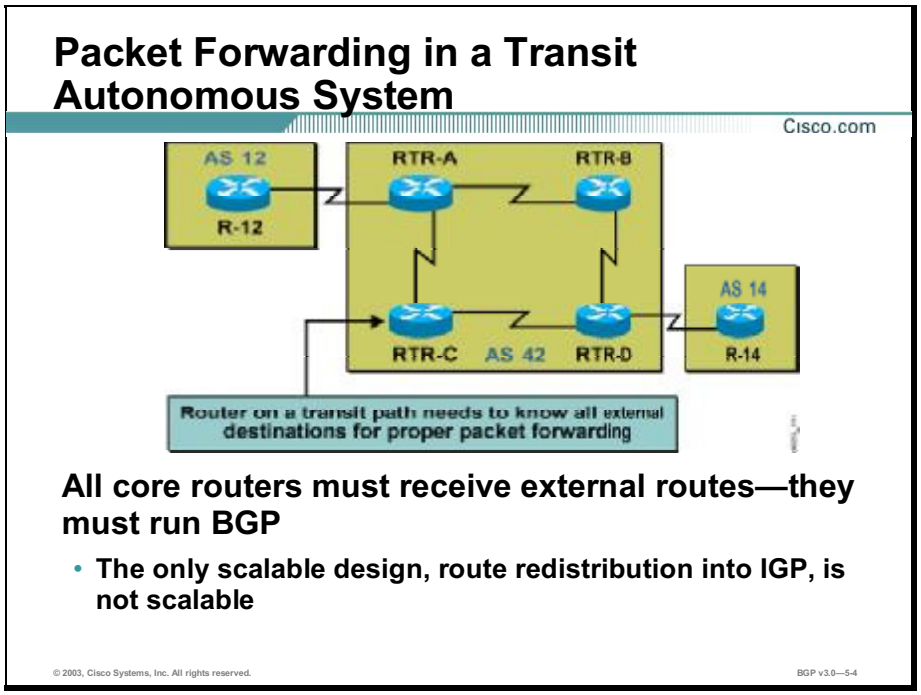
This lesson includes these topics:

- Overview
- Packet Forwarding in a Transit Autonomous System
- Recursive Lookup in Cisco IOS Software
- Routing Protocols in a Transit Autonomous System
- BGP and IGP Interaction
- Problems with BGP and IGP Interaction
- Summary
- Assessment (Quiz): Packet Forwarding in Transit Autonomous Systems

 Copyright © 2003, Cisco Systems, Inc.

Packet Forwarding in a Transit Autonomous System

This topic describes packet forwarding in a transit AS.


Packet Forwarding in a Transit Autonomous System
Cisco.com
AS 12
R-12
RTR-A
RTR-B
RTR-C
AS 42
RTR-D
AS 14
R-14
Router on a transit path needs to know all external destinations for proper packet forwarding
All core routers must receive external routes—they must run BGP
• The only scalable design, route redistribution into IGP, is not scalable
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-4

When BGP updates have propagated through the transit AS to all neighboring autonomous systems, the IP traffic can start to flow.

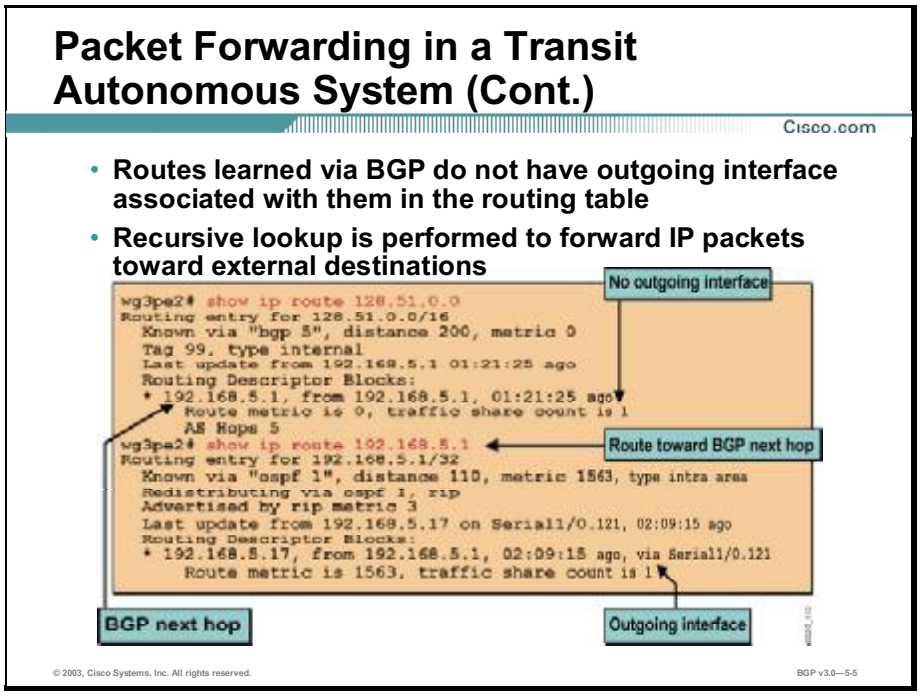
Router R-14 will forward to RTR-D IP packets with the destination address matching a network in AS 12. RTR-D will check its routing table and find that there is a BGP route for that destination. The BGP route has a next-hop reference, which points to the far end of the EBGP session between R-12 and RTR-A. So RTR-D once again checks the routing table and finds that it should forward the packet to RTR-C in this case.

Thus, RTR-C receives the IP packet with a destination address indicating a host within AS 12. In order to be able to forward this packet, RTR-C must have a matching route in its routing table. A default route or gateway of last resort is not appropriate because in the next instant RTR-C could receive another packet, coming in from the other direction and destined to AS 14.

The conclusion is that both RTR-C and RTR-B, to handle all possible cases, must have routing information to all the external networks that RTR-A and RTR-D have. The only scalable way of providing routers with this information is to update RTR-C and RTR-B with IBGP from both RTR-A and RTR-D.

In theory, the external information received by RTR-A and RTR-D, respectively, could be redistributed by these ingress routers into the IGP in use within the transit AS. However, no IGP can handle the volume of information that BGP can. So there would always be a risk that the IGP would break due to information overload, causing a total network meltdown in the AS. The volume of routing information carried by BGP in the contemporary Internet has long ago passed the limits for what it is possible to carry in any IGP.

 Copyright © 2003, Cisco Systems, Inc.


Packet Forwarding in a Transit Autonomous System (Cont.)
Cisco.com
• Routes learned via BGP do not have outgoing interface associated with them in the routing table
• Recursive lookup is performed to forward IP packets toward external destinations
wg3pe2# show ip route 128.51.0.0
Routing entry for 128.51.0.0/16
Known via "bgp 5", distance 200, metric 0
Tag 99, type internal
Last update from 192.168.5.1 01:21:25 ago
Routing Descriptor Blocks:
* 192.168.5.1, from 192.168.5.1, 01:21:25 ago
Route metric is 0, traffic share count is 1
AS Hops 5
wg3pe2# show ip route 192.168.5.1
Routing entry for 192.168.5.1/32
Known via "ospf 1", distance 110, metric 1563, type intra area
Redistributing via ospf 1, rip
Advertised by rip metric 3
Last update from 192.168.5.17 on Serial1/0.121, 02:09:15 ago
Routing Descriptor Blocks:
* 192.168.5.17, from 192.168.5.1, 02:09:15 ago, via Serial1/0.121
Route metric is 1563, traffic share count is 1
No outgoing interface
Route toward BGP next hop
BGP next hop
Outgoing interface
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-5

A BGP route is installed in a router's IP routing table only if the IP address in the next-hop attribute is reachable according to the information already in the routing table. The installed BGP route contains a reference to that next-hop address. So, the network will be reachable via an IP address, which may or may not be directly connected. Because there is no clear reference to a physical interface, the BGP route is installed in the IP routing table without any information about outgoing interface.

The router must evaluate the recursive reference to the BGP next hop sooner or later in order to allow packet forwarding toward external destinations. The point in time when the recursive reference is resolved is dependent upon the IP switching mechanism used by the router. At the latest, the router performs the recursive lookup when an IP packet with a destination address matching the BGP route should be forwarded. The router determines which outgoing interface should be used and which Layer 2 address to assign (if applicable). The router creates a cache entry so that successive IP packets to the same destination can be routed using the same outgoing interface and Layer 2 address.

Practice

Q1) Why do you need to run IBGP on all core routers?

A) to ensure that a full mesh exists between all BGP routers in the AS

B) to allow routers to properly resolve the next-hop address

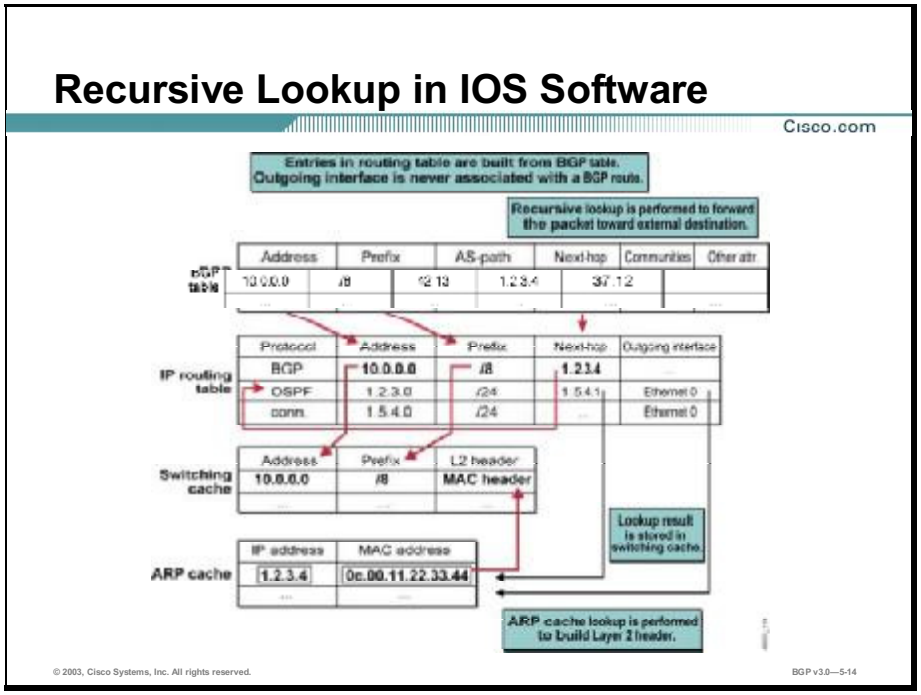
C) to allow routers to forward packets toward all external destinations

D) to ensure correct propagation of the gateway of last resort

 Copyright © 2003, Cisco Systems, Inc.

Recursive Lookup in Cisco IOS Software

This topic explains how recursive lookup functions in Cisco IOS software.



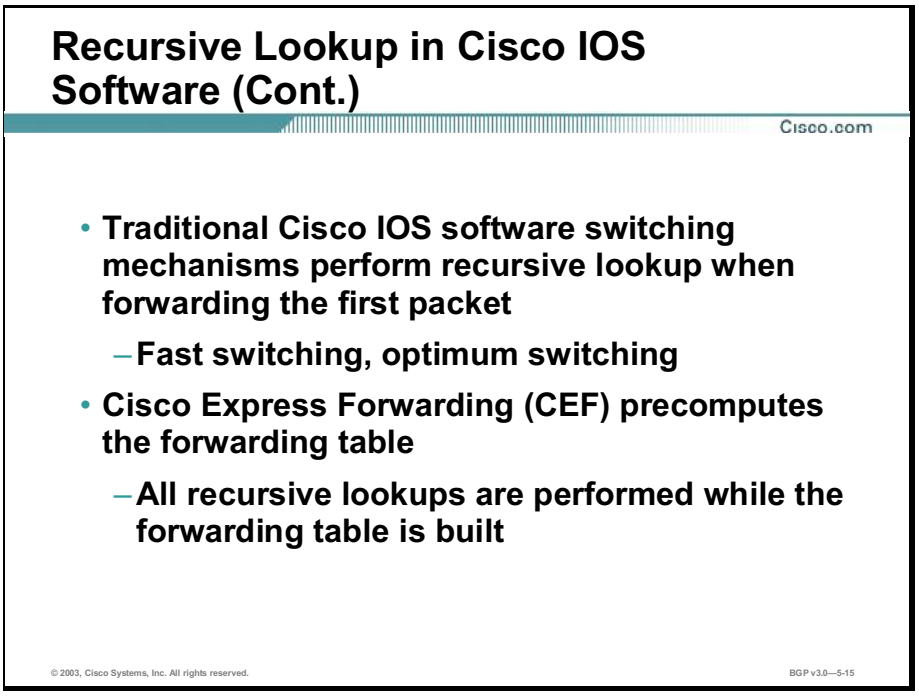
The figure here presents the steps in the recursive lookup process in Cisco IOS software. The router has received a BGP update about network 10.0.0.0/8. It was associated with an AS-path attribute set to 42 13, a next-hop attribute set to the IP address 1.2.3.4, and a community value 37:12. Some other attributes were also carried with the update.

Because the next-hop address 1.2.3.4 is reachable according to the routing table, the BGP route is also installed in the routing table. Network number, subnet mask, and next-hop attributes are inherited from the BGP table. No outgoing interface is assigned.

When an IP packet with a destination in network 10.0.0.0 is received, the router searches the routing table and finds the installed BGP route. The router takes the indicated next-hop address 1.2.3.4 and searches the routing table again. It will now find a match with the Open Shortest Path First (OSPF) route to subnet 1.2.3.0/24. The 1.2.3.0/24 route has an outgoing interface set to interface ethernet 0 and a next hop set to 1.5.4.1, meaning that packets destined for network 10.0.0.0 should be forwarded via 1.5.4.1, which is directly reachable over ethernet 0.

The Address Resolution Protocol (ARP) table is used to find the MAC address for IP address 1.5.4.1. The MAC address is used to forward the IP packet to network 10.0.0.0 out the ethernet 0 interface. The MAC header is stored in the cache for successive packets to network 10.0.0.0.

Note This example illustrates the recursive lookup performed when the router uses cache-based IP switching mechanisms; for example, fast switching or optimum switching. See the next figure for more information on the differences between Cisco Express Forwarding (CEF) and the cache-based switching mechanisms.


Recursive Lookup in Cisco IOS Software (Cont.)
Cisco.com
• Traditional Cisco IOS software switching mechanisms perform recursive lookup when forwarding the first packet
– Fast switching, optimum switching
• Cisco Express Forwarding (CEF) precomputes the forwarding table
– All recursive lookups are performed while the forwarding table is built
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-15

Traditional Cisco IOS switching mechanisms used the traffic-driven, cache-based switching approach. Both fast switching and optimum switching populate the IP switching cache on demand, meaning that before any IP packets are forwarded, the cache is empty. After the first packet to a specific destination arrives, all routing table lookups are being done, including recursive lookup in the case of a BGP route. The result of the lookup is being cached for later use when successive packets to the same destination arrive. The process is repeated for every specific destination.

CEF prebuilds a complete IP forwarding table (called the Forwarding Information Base [FIB]) based on the IP routing table. After the router installs a routing entry into its routing table, incoming routing information updates trigger the recursive lookup, and the outgoing interface and the actual physical next hop of the route are determined. MAC address resolution and MAC header generation are still traffic-driven and stored in the cache.

Practice

Q1) What is recursive lookup in Cisco IOS software?

A) verifying that the route exists both in the BGP table and the IP routing table

B) resolving the MAC address of the route to build the MAC header

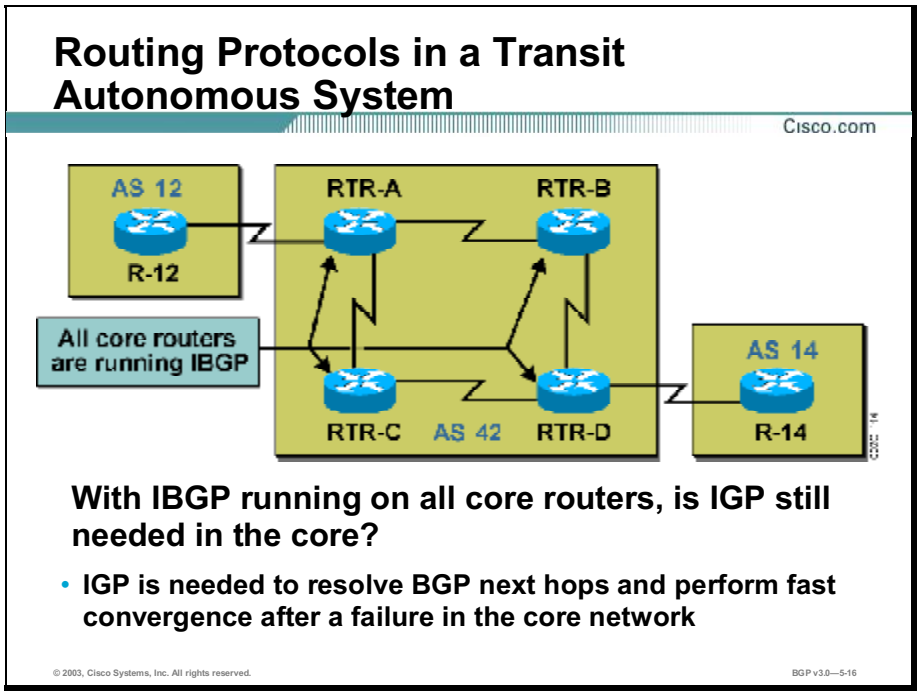
C) searching the routing table to find a path to the next-hop address

D) the process used to build the IP fast cache on the Cisco IOS router

 Copyright © 2003, Cisco Systems, Inc.

Routing Protocols in a Transit Autonomous System

This topic explains the need for an IGP in a transit backbone running BGP on all routers.


Routing Protocols in a Transit Autonomous System
Cisco.com
AS 12
R-12
RTR-A
RTR-B
All core routers are running IBGP
RTR-C
AS 42
RTR-D
AS 14
R-14
With IBGP running on all core routers, is IGP still needed in the core?
• IGP is needed to resolve BGP next hops and perform fast convergence after a failure in the core network
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-16

Some network designers base their network design on the wrong assumption that an internal routing protocol is not needed in a transit AS where all routers run BGP. However, the internal routing protocol is still needed inside an AS for two reasons:

- To provide routing information needed to establish the IBGP sessions
- To resolve next-hop references (recursive routing)

For example, when RTR-D in the example here receives an IP packet with the destination in AS 12, it will do a recursive lookup to find the outgoing interface to be used for packet forwarding. It performs the recursive lookup based on IGP information. If there is suddenly an internal problem within AS 42, and the next-hop address is reachable a different way, the IGP will determine this fact. The IGP route to the next-hop network will be changed by the router due to newly received IGP route information, and all cache entries relying on the old information will be invalidated. The next recursive lookup that RTR-D performs will now indicate a different outgoing interface than before the problem occurred.

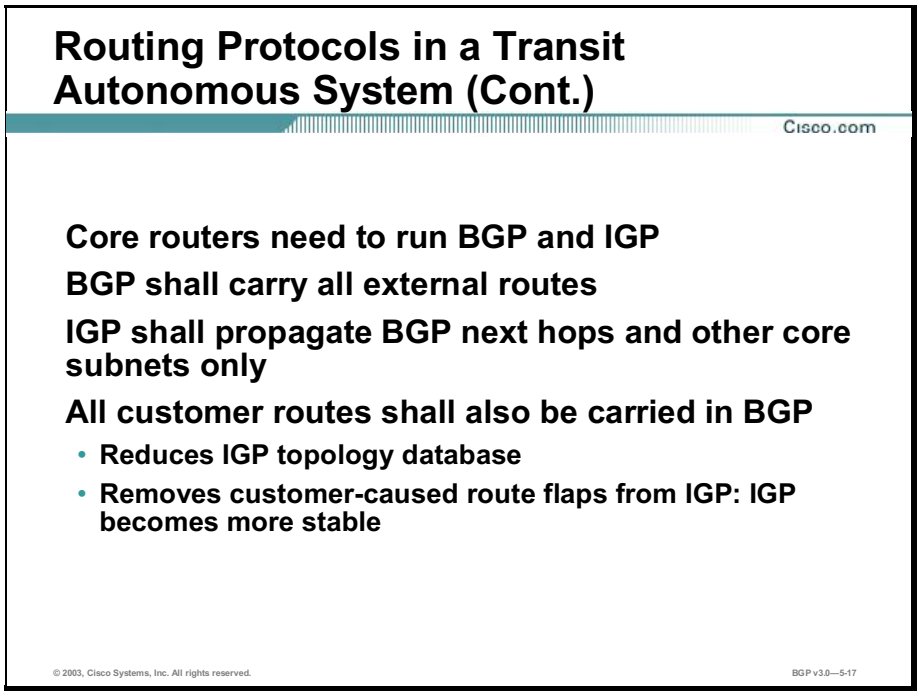
During the IGP convergence process, the BGP routing is not affected. The only routing update exchanged during the transit AS convergence would be IGP updates describing how to reach internal destinations (including the far ends of the EBGP sessions).

The packet forwarding to external destinations thus benefits from the high-speed convergence offered by the IGP. The faster the IGP determines that it should use an alternate path within the

AS to reach the next-hop address, the faster it will re-establish IP connectivity toward external destinations.

The conclusion is that an IGP is still needed inside a transit AS, and the network will work better if it is an IGP with fast convergence.

 Copyright © 2003, Cisco Systems, Inc.


Routing Protocols in a Transit Autonomous System (Cont.)
Cisco.com
Core routers need to run BGP and IGP
BGP shall carry all external routes
IGP shall propagate BGP next hops and other core subnets only
All customer routes shall also be carried in BGP
• Reduces IGP topology database
• Removes customer-caused route flaps from IGP: IGP becomes more stable
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-17

Both BGP and the configured IGP should be configured on all core routers inside the transit AS. The IGP should carry as little information as possible—ideally only the links within the core network, the loopback interfaces, and the external subnets used in EBGP sessions with neighboring autonomous systems. This information is enough to establish IBGP sessions and resolve next-hop addresses. The IGP will also work better if it carries less routing information.

No routes external to the transit AS should ever be redistributed by any router from BGP into the IGP. All external routes should be in BGP only.

In autonomous systems that provide customer connectivity (not only transit service), it is also highly recommended that the customer networks be carried in BGP in order to reduce the amount of information in the IGP and increase IGP stability.

Practice

Q1) What are two reasons why you need an IGP in a transit AS? (Choose two.)

A) to provide the gateway of last resort to BGP speakers

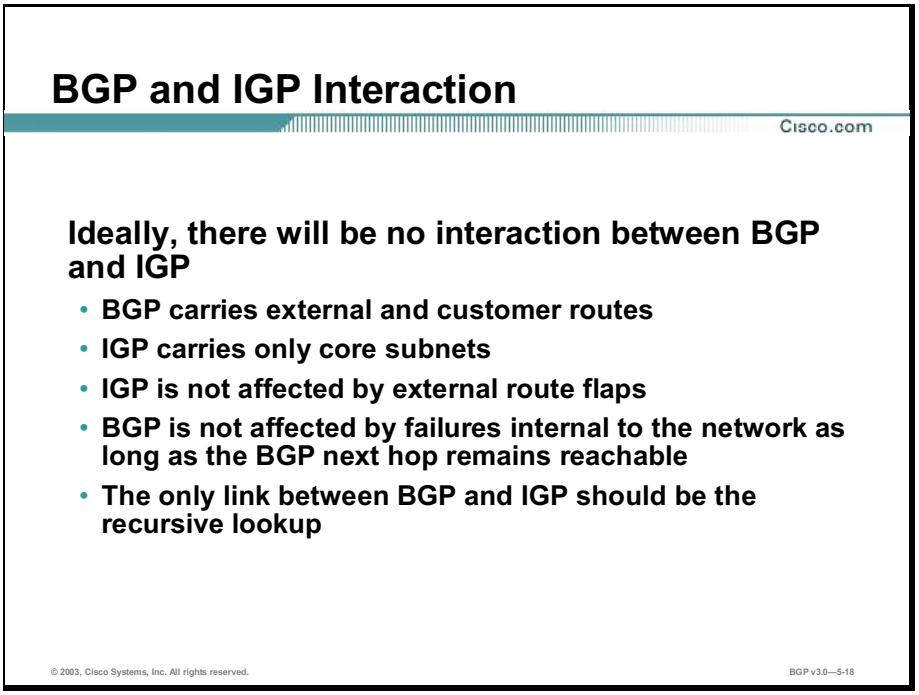
B) to perform the recursive lookup for external networks

C) to carry BGP routes through the transit AS

D) to establish IBGP sessions between nondirectly connected routers

BGP and IGP Interaction

This topic describes the interaction between BGP and IGPs in a transit AS.


BGP and IGP Interaction
Cisco.com
Ideally, there will be no interaction between BGP and IGP
• BGP carries external and customer routes
• IGP carries only core subnets
• IGP is not affected by external route flaps
• BGP is not affected by failures internal to the network as long as the BGP next hop remains reachable
• The only link between BGP and IGP should be the recursive lookup
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-18

Ideally, BGP and the IGP carry two different sets of routing information. BGP carries those routes received from other autonomous systems and those routes that belong to the local AS and should be announced to other autonomous systems. The IGP carries only enough information to establish IBGP sessions and resolve the next-hop addresses.

The IGP will provide reachability toward the BGP next-hop addresses only if it is not disturbed by external updates from other autonomous systems.

BGP should take care of the external information. As long as the IGP finds a usable way to the BGP next hops, the BGP does not need to do any recalculation due to internal problems within the AS.

 Copyright © 2003, Cisco Systems, Inc.

BGP and IGP Interaction (Cont.)

Cisco.com

Sometimes, BGP and IGP will propagate the same route

- Usually due to bad network design
- In this case, routes are determinedd in EBGP/IGP/IBGP order based on administrative distances of the routes

Routing protocol	Default administrative distance
EBGP	20
IGP	90 - 170
IBGP	200

© 2003, Cisco Systems, Inc. All rights reserved.

Sometimes the interaction between BGP and the IGP is not ideal due to a number of reasons, including bad network design. In the worst case, the same networks might be carried in both the IGP and in BGP. For example, the subnets connecting the AS with neighboring autonomous systems have to be announced via IGP to enable next-hop resolution but may also be announced via BGP by the remote AS or the local AS. In any case, information about the same IP prefix will appear in both the IGP and the BGP data structures.

When the router installs routing information into the routing table, it checks to see whether there are several sources of information for a particular IP prefix. If so, the router will install the information that it determines is most reliable. The administrative distance (AD) determines which source to use.

BGP considers both EBGP and IBGP routes in the BGP selection process. BGP will therefore never try to install both an EBGP route and an IBGP route for the same destination. Comparison between ADs will thus occur only when two different protocols carry the same destination network.

If BGP selects an EBGP route as the best route for a given destination network, it will try to install that route with a very low AD, meaning that routes learned via EBGP have a high likelihood of being installed in the forwarding table.

If BGP selects an IBGP route as the best, it will try to install it with a high AD, meaning that routes learned via IBGP have a low likelihood of being installed in the forwarding table.

All IGPs, such as Enhanced Interior Gateway Routing Protocol (EIGRP), OSPF, Intermediate System-to-Intermediate System (IS-IS), and so on, have a medium likelihood of being installed. The ADs for IGPs fall between the ADs of EBGP and IBGP.

Note	The reason for giving EBGP a low default AD is because EBGP indicates routes external to the local AS. IP packets with destination addresses to those networks should leave the AS sooner rather than later. It is, in most cases, better that they leave the AS right away.

Practice

Q1) Why should you transport your customer routes in BGP and not in an IGP?

A) to ensure that the gateway of last resort propagates across the AS

B) to correctly resolve addresses in the next-hop attribute

C) to ensure full reachability to external network destinations

D) to protect the IGP from carrying too many routes

Q2) How should BGP react to a failure inside a transit AS that has redundant paths?

A) BGP sessions across the failed link will terminate.

B) BGP does nothing because the IGP finds an alternate path to the neighbors.

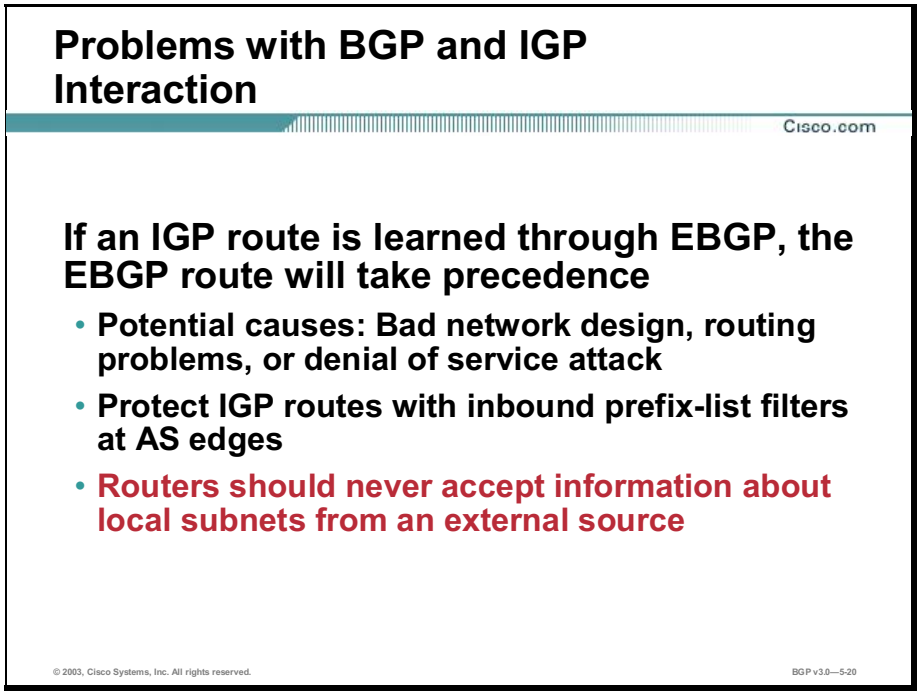
C) Failure within the AS causes the BGP to recalculate new paths.

D) BGP will terminate the link between itself and IGP.

 Copyright © 2003, Cisco Systems, Inc.

Problems with BGP and IGP Interaction

This topic explains the potential problems that might arise from BGP and IGP interaction.


Problems with BGP and IGP Interaction
Cisco.com
If an IGP route is learned through EBGP, the EBGP route will take precedence
• Potential causes: Bad network design, routing problems, or denial of service attack
• Protect IGP routes with inbound prefix-list filters at AS edges
• Routers should never accept information about local subnets from an external source
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-20

If routing information about the same IP prefix is learned via both EBGP and an IGP, the router will use the EBGP information. If an external AS is feeding the local AS with EBGP routes that actually should be local, routers within the AS will erroneously forward IP packets destined to those local networks out of the local AS.

There are several potential reasons for this behavior; the most common is that the remote AS is improperly configured or there is a denial of service (DoS) attack. To protect a local AS from this undesired behavior, network administrators should install inbound filters on all EBGP sessions to filter incoming routes and reject routing information about networks that are actually local to the AS.

Practice

Q1) What happens when the same route is learned via BGP and an IGP?

A) BGP will install both an EBGP and an IBGP route for the same destination.

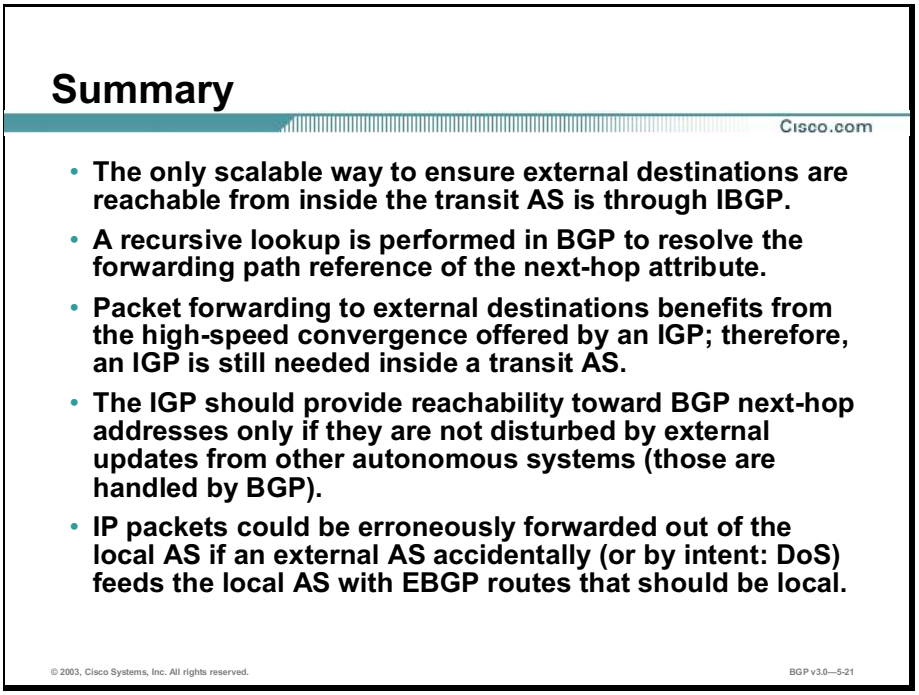
B) If the same route is learned via both the IGP and EBGP, the EBGP route is preferred.

C) The route learned through the IGP will be installed in the routing table.

D) If the same route is learned via both the IGP and IBGP, the IBGP route is preferred.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• The only scalable way to ensure external destinations are reachable from inside the transit AS is through IBGP.
• A recursive lookup is performed in BGP to resolve the forwarding path reference of the next-hop attribute.
• Packet forwarding to external destinations benefits from the high-speed convergence offered by an IGP; therefore, an IGP is still needed inside a transit AS.
• The IGP should provide reachability toward BGP next-hop addresses only if they are not disturbed by external updates from other autonomous systems (those are handled by BGP).
• IP packets could be erroneously forwarded out of the local AS if an external AS accidentally (or by intent: DoS) feeds the local AS with EBGP routes that should be local.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-21

Next Steps

After completing this lesson, go to:

- Configuring a Transit Autonomous System lesson

References

For additional information, refer to these resources:

- For more information on transit autonomous systems, refer to “BGP Case Studies Section 1” at the following URL: http://www.cisco.com/warp/public/459/13.html

 Copyright © 2003, Cisco Systems, Inc.

Quiz: Packet Forwarding in Transit Autonomous Systems

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe packet forwarding in a transit AS
- Explain recursive lookup in Cisco IOS software
- Explain the need for an IGP in a transit backbone running BGP on all routers
- Describe interactions between BGP and IGP
- Explain the potential problems that might arise from BGP and IGP interaction

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What are two reasons why you must run IBGP on all routers within a transit backbone? (Choose two.)

A) so routers can properly forward packets toward all external destinations

B) to ensure that a full mesh exists between all routers in the AS

C) to allow routers to properly process the BGP next-hop attribute

D) because IGPs cannot scale large enough to handle redistribution of BGP routes

Copyright © 2003, Cisco Systems, Inc.

Q2) If a transit backbone has IBGP running on all routers, what are two reasons why it is still necessary to use an IGP? (Choose two.)

A) to provide routing information needed to establish the IBGP sessions

B) to resolve next-hop references used in recursive routing

C) so that BGP routes can be properly transported through the AS

D) to provide user workstations with a network default gateway

Q3) What is the AD of the following protocols? (Fill in the blanks.)

A) IBGP _____

B) EBGP _____

C) OSPF _____

D) IS-IS _____

E) RIP _____

Q4) What are two reasons why the AD is an important consideration for BGP network design? (Choose two.)

A) The AD affects how routes are selected for use in the IP routing table.

B) The AD controls how routing information is entered into the BGP table.

C) If a route is advertised by both an IGP and through EBGP, the router will prefer the external route.

D) AD is not a large concern to BGP design, because the router will always choose the route advertised by the protocol best suited to reach the destination.

 Copyright © 2003, Cisco Systems, Inc.

Q5) With regard to recursive route lookups, what are two ways in which CEF is different from traditional Cisco IOS switching mechanisms such as route caching? (Choose two.)

A) Traditional Cisco IOS switching mechanisms wait for the first packet to arrive before recursive lookup can take place.

B) New entries in the IP routing table will trigger a recursive lookup in traditional Cisco IOS switching mechanisms.

C) CEF prebuilds a complete IP forwarding table based on the IP routing table.

D) CEF will build a FIB directly from the entries in the BGP table prior to any BGP packets arriving at the router.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Configuring a Transit Autonomous System

Overview

Specifying an autonomous system (AS) as a transit backbone introduces specific requirements in the design, scaling, and configuration of Border Gateway Protocol (BGP). This lesson introduces the configuration requirements of Internal Border Gateway Protocol (IBGP) to implement a transit AS. Configuration details of IBGP are discussed in this lesson, including IBGP neighbor configuration, using loopback interfaces for IBGP neighbors, disabling BGP synchronization, and modifying the default administrative distances (ADs) of BGP. This lesson concludes with a discussion of the scalability concerns of BGP in the transit backbone.

Importance

Configuring a BGP network in a transit services configuration requires special care to ensure consistency of routing information throughout the AS. Understanding the configuration requirements for a transit backbone is crucial to successfully implementing the transit autonomous network.

Objectives

Upon completing this lesson, you will be able to:

- Identify the Cisco IOS® commands required to configure IBGP neighbors
- Identify the Cisco IOS commands required to configure IBGP sessions between loopback interfaces
- Identify the Cisco IOS commands required to configure BGP synchronization to ensure successful IBGP operation of the transit AS

- Identify the Cisco IOS commands required to change the AD of BGP routes
- Identify the scalability limitations of IBGP-based backbones

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

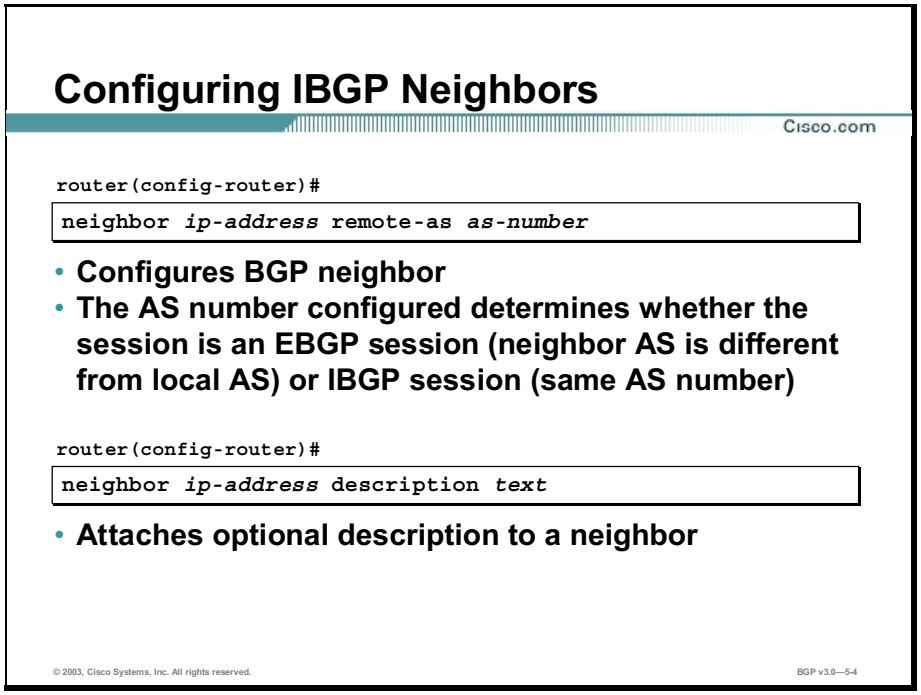
This lesson includes these topics:

- Overview
- Configuring IBGP Neighbors
- Configuring IBGP Sessions Between Loopback Interfaces
- Configuring BGP Synchronization
- Changing the Administrative Distance of BGP Routes
- Scalability Limitations of IBGP-Based Transit Backbones
- Summary
- Assessment (Quiz): Configuring a Transit Autonomous System

 Copyright © 2003, Cisco Systems, Inc.

Configuring IBGP Neighbors

This topic lists the Cisco IOS commands required to configure IBGP neighbors in an AS.


Configuring IBGP Neighbors
Cisco.com
router(config-router)#
neighbor ip-address remote-as as-number
• Configures BGP neighbor
• The AS number configured determines whether the session is an EBGP session (neighbor AS is different from local AS) or IBGP session (same AS number)
router(config-router)#
neighbor ip-address description text
• Attaches optional description to a neighbor
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-4

Configuring an IBGP neighbor is a simple task:

- Configure a BGP neighbor, specifying the same AS number.
- Optionally, attach a description to the neighbor to help in documentation and troubleshooting efforts.

neighbor remote-as

To add an entry to the BGP neighbor table, use the **neighbor remote-as** router configuration command.

neighbor [*ip-address* | *peer-group-name*] **remote-as** *as-number*

To remove an entry from the table, use the **no** form of this command.

no neighbor [*ip-address* | *peer-group-name*] **remote-as** *as-number*

Syntax Description

ip-address	Neighbor IP address
peer-group-name	Name of a BGP peer group
as-number	AS to which the neighbor belongs

neighbor description

To associate a description with a neighbor, use the **neighbor description** router configuration command.

neighbor [*ip-address* | *peer-group-name*]**description** *text*

To remove the description, use the **no** form of this command.

no neighbor [*ip-address* | *peer-group-name*] **description** *text*

Syntax Description

ip-address	Neighbor IP address
peer-group-name	Name of a BGP peer group
text	Text (up to 80 characters) that describes the neighbor

Practice

Q1) Which Cisco IOS command do you use to configure a description on a BGP session?

A) **neighbor** *ip-address* **remote-as** *number* **description** *text*

B) **description** *text*

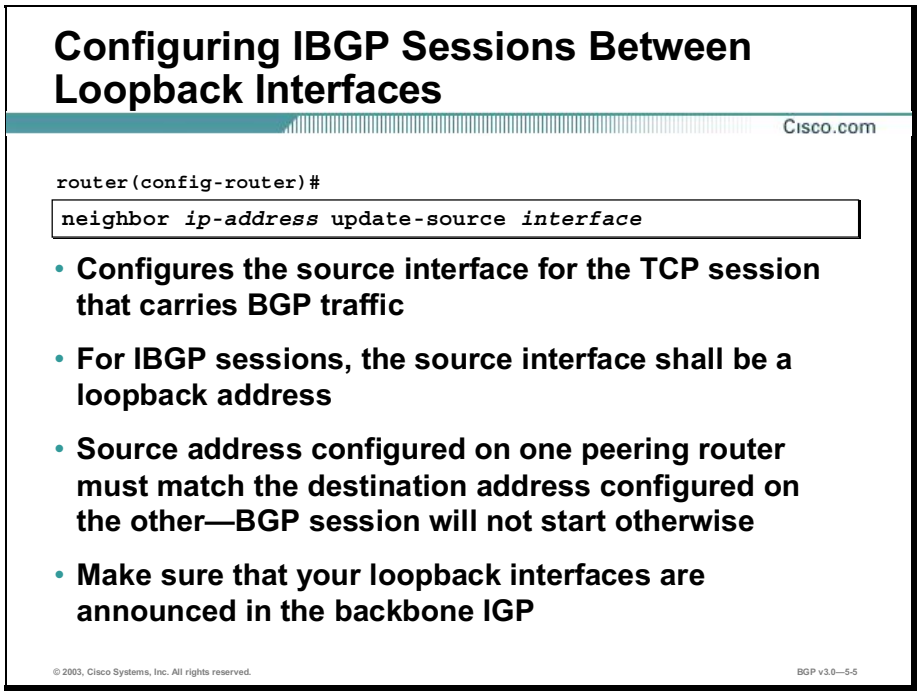
C) **neighbor** *ip-address* **description** *text*

D) **ip bgp description** *text*

 Copyright © 2003, Cisco Systems, Inc.

Configuring IBGP Sessions Between Loopback Interfaces

This topic presents the Cisco IOS command required to configure IBGP sessions between loopback interfaces on routers in a common AS.


Configuring IBGP Sessions Between Loopback Interfaces
Cisco.com
router(config-router)#
neighbor ip-address update-source interface
• Configures the source interface for the TCP session that carries BGP traffic
• For IBGP sessions, the source interface shall be a loopback address
• Source address configured on one peering router must match the destination address configured on the other—BGP session will not start otherwise
• Make sure that your loopback interfaces are announced in the backbone IGP
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-5

When a BGP session is established between two routers, both routers attempt to set up the TCP connection by sending TCP SYN packets to each other. If both succeed, one of the sessions will be brought down so only one remains. The TCP packets will have the destination IP address as configured with the **neighbor** command. But they must also have a source IP address assigned. If no update source is configured, the router will set the source IP address of the outgoing TCP session to the IP address of the outgoing physical interface.

When a TCP SYN packet with the BGP well-known port number arrives at the peer router, the receiver checks if the connection attempt is coming from one of the configured peers. If the source IP address is not in the list of configured neighbors, the receiver denies the connection attempt.

As a general rule, IBGP sessions should be established between loopback interfaces of BGP-speaking routers. The destination IP address configured in the **neighbor** statement should therefore be the IP address of the loopback interface of the peer router. But the local router must also make sure that the source address of the outgoing TCP connection attempt is the IP address that the peer router has listed. Configuring BGP neighbors using **update-source** ensures that the source address of the outgoing TCP connection is correct by referring to the interface that has the correct IP address. Normally this interface is the loopback interface of the local router.

neighbor update-source

To instruct Cisco IOS software to allow IBGP sessions to use any operational interface for TCP connections, use the **neighbor update-source** router configuration command.

neighbor [*ip-address* | *peer-group-name*] **update-source** *interface*

To restore the interface assignment to the closest interface, which is called the "best local address," use the **no** form of this command.

no neighbor [*ip-address* | *peer-group-name*] **update-source** *interface*

Syntax Description

ip-address	Neighbor IP address
peer-group-name	Name of a BGP peer group
interface	Loopback interface

Practice

Q1) Which Cisco IOS command do you use to configure a BGP session between loopback interfaces?

A) **neighbor** *ip-address* **remote-as** *number* **update-source** *interface*

B) **ip bgp source-interface** *interface*

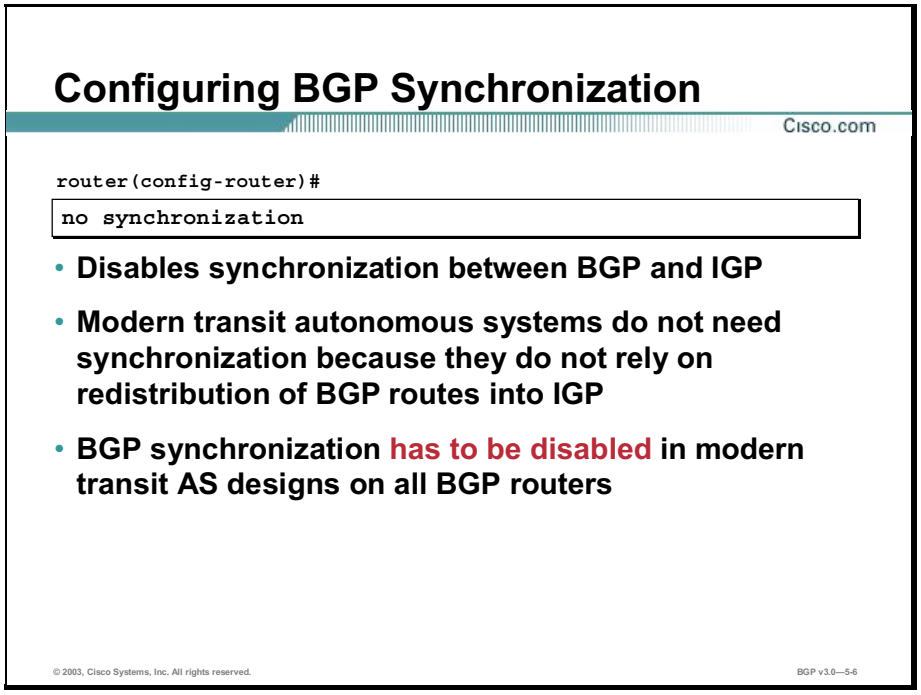
C) **neighbor** *ip-address* **update-source** *interface*

D) **ip bgp update-source** *interface*

 Copyright © 2003, Cisco Systems, Inc.

Configuring BGP Synchronization

This topic presents the Cisco IOS command required to configure BGP synchronization to ensure successful IBGP operation of a transit AS.


Configuring BGP Synchronization
Cisco.com
router(config-router)#
no synchronization
• Disables synchronization between BGP and IGP
• Modern transit autonomous systems do not need synchronization because they do not rely on redistribution of BGP routes into IGP
• BGP synchronization has to be disabled in modern transit AS designs on all BGP routers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-6

The BGP synchronization rule states that if an AS provides transit service to another AS, BGP should not advertise a route until all of the routers within the AS have learned about the route via an IGP. Network designers used synchronization in older transit AS designs that relied on BGP route redistribution into the IGP. Modern AS designs do not rely on this feature anymore because the number of routes carried in the Internet exceeds the scalability range of any known IGP. Redistribution into IGP is thus no longer applicable, and you must disable the synchronization feature for your transit AS to work.

synchronization

To enable the Cisco IOS software to advertise a network route without waiting for the IGP (disable synchronization between BGP and your IGP), use the **no** form of the **synchronization** command.

no synchronization

Syntax Description

This command has no arguments or keywords.

Practice

Q1) What BGP parameter do you need to disable for proper IBGP operation?

A) redistribution

B) AS-path propagation

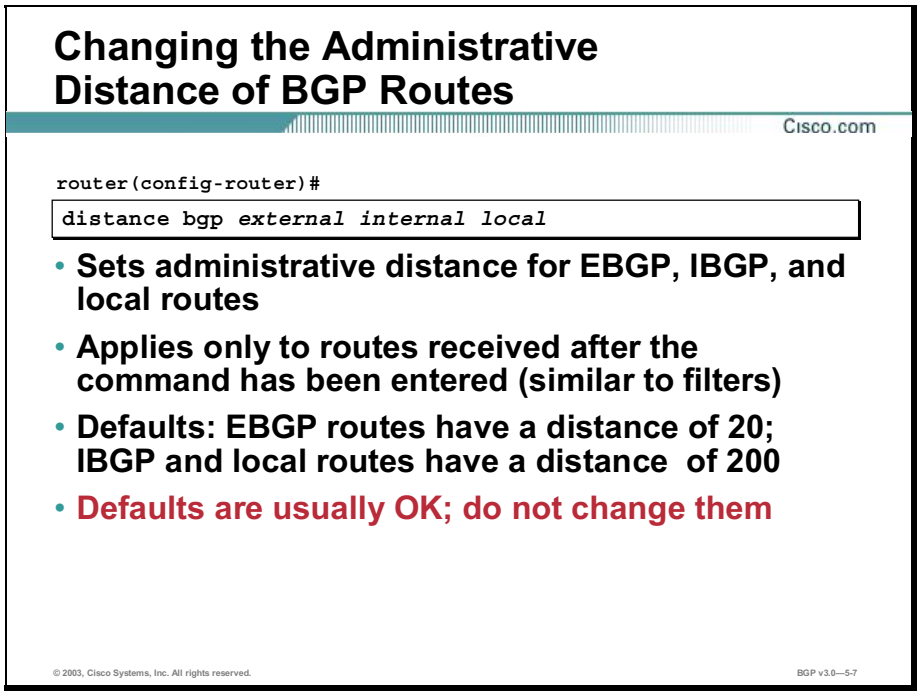
C) synchronization

D) split horizon

 Copyright © 2003, Cisco Systems, Inc.

Changing the Administrative Distance of BGP Routes

This topic presents the Cisco IOS command required to change the AD of BGP routes.


Changing the Administrative Distance of BGP Routes
Cisco.com
router(config-router)#
distance bgp external internal local
• Sets administrative distance for EBGP, IBGP, and local routes
• Applies only to routes received after the command has been entered (similar to filters)
• Defaults: EBGP routes have a distance of 20; IBGP and local routes have a distance of 200
• Defaults are usually OK; do not change them
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-7

distance bgp

To allow the use of external, internal, and local ADs that could be a better route to a node, use the **distance bgp** router configuration command.

distance bgp *external-distance internal-distance local-distance*

To return to the default values, use the **no** form of this command.

no distance bgp

Syntax Description

external-distance	AD for BGP external routes. External routes are routes for which the best path is learned from a neighbor external to the AS. Acceptable values are from 1 to 255. The default is 20. Routes with a distance of 255 are not installed in the routing table.
internal-distance	AD for BGP internal routes. Internal routes are those routes that are learned from another BGP entity within the same AS. Acceptable values are from 1 to 255. The default is 200. Routes with a distance of 255 are not installed in the routing table

local-distance	AD for BGP local routes. Local routes are those networks listed with a network router configuration command, often as backdoors (BGP backdoor makes the IGP route the preferred route) for that router or for networks that are being redistributed from another protocol. Acceptable values are from 1 to 255. The default is 200. Routes with a distance of 255 are not installed in the routing table

Practice

Q1) How can you change the AD of BGP routes?

A) through AS-path prepending

B) with the **distance bgp** router configuration command

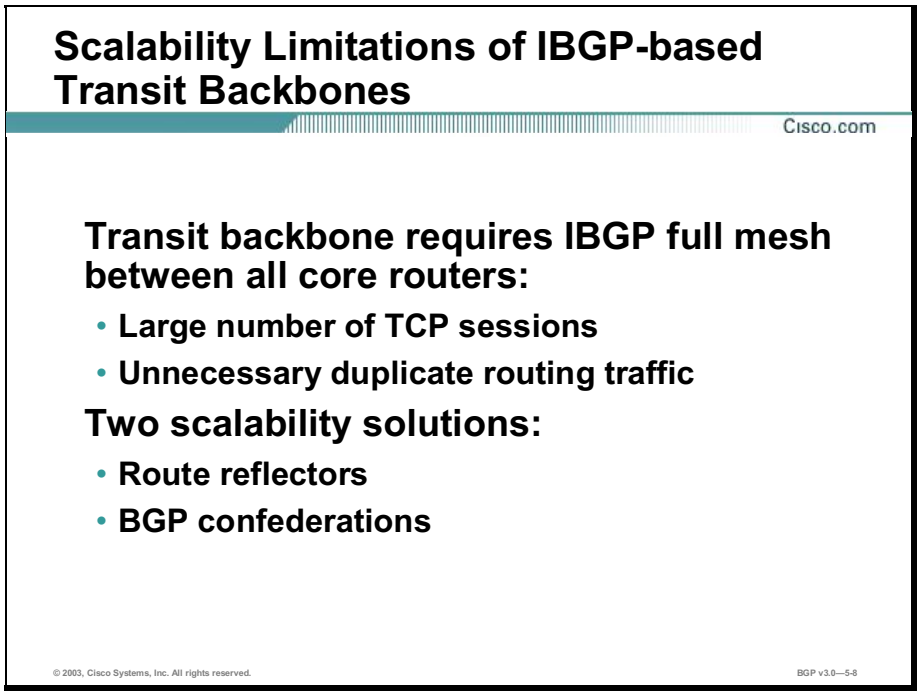
C) by redistributing BGP routes into an IGP

D) by disabling BGP synchronization on the router

 Copyright © 2003, Cisco Systems, Inc.

Scalability Limitations of IBGP-Based Transit Backbones

This topic identifies the scalability limitations of IBGP-based backbones.


Scalability Limitations of IBGP-based Transit Backbones
Cisco.com
Transit backbone requires IBGP full mesh between all core routers:
• Large number of TCP sessions
• Unnecessary duplicate routing traffic
Two scalability solutions:
• Route reflectors
• BGP confederations
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-8

IBGP split-horizon rules, as documented in previous lessons, mandate an IBGP connection between every border router and every other BGP router in an AS.

The general design rule in IBGP design is to have a full mesh of IBGP sessions. But, a full mesh of IBGP session between “n” number of routers would require (n * (n-1)) / 2 IBGP sessions. For example, a full mesh between 10 routers would require (10 * 9) / 2 = 45 IBGP sessions.

Since every IBGP session on a router uses a separate TCP session, an update that must be sent by the router to all IBGP peers must be sent on each of the TCP sessions. If a router is attached to the rest of the network over just a single link, this single link has to carry all TCP/IP packets for all IBGP sessions. This situation results in duplication of the update over the single link.

Two different solutions are available:

- The route reflector solution modifies the IBGP split-horizon rules and allows a particular router to forward (under certain conditions) incoming IBGP updates to a select group of IBGP neighbors. The router performing this function is the “route reflector.”

- The BGP confederations solution introduces the concept of a number of smaller autonomous systems within the original AS. These smaller autonomous systems exchange BGP updates between themselves using intraconfederation EBGP sessions.

Practice

Q1) What are the scalability limitations of an IBGP-based transit AS?

A) The transit backbone can have only two exit points.

B) All EBGP speakers must have a BGP session to all BGP-speaking routers in the transit AS.

C) The transit area is limited to a maximum distance of 100 routers.

D) A full mesh of IBGP sessions is required due to BGP split-horizon rules.

Q2) What two BGP tools can you use to overcome IBGP scalability issues in a transit AS? (Choose two.)

A) disabling synchronization

B) route reflectors

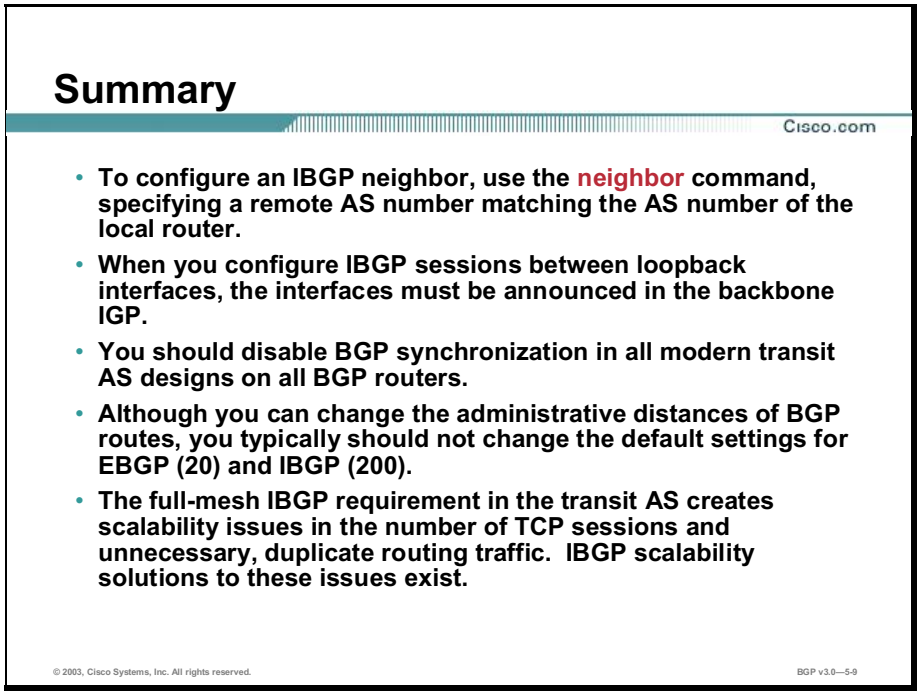
C) BGP confederations

D) BGP peer groups

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• To configure an IBGP neighbor, use the neighbor command, specifying a remote AS number matching the AS number of the local router.
• When you configure IBGP sessions between loopback interfaces, the interfaces must be announced in the backbone IGP.
• You should disable BGP synchronization in all modern transit AS designs on all BGP routers.
• Although you can change the administrative distances of BGP routes, you typically should not change the default settings for EBGP (20) and IBGP (200).
• The full-mesh IBGP requirement in the transit AS creates scalability issues in the number of TCP sessions and unnecessary, duplicate routing traffic. IBGP scalability solutions to these issues exist.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-9

Next Steps

After completing this lesson, go to:

- Monitoring and Troubleshooting IBGP in Transit Autonomous Systems lesson

References

For additional information, refer to these resources:

- For more information on transit autonomous systems, refer to “BGP Case Studies Section 1” at the following URL: http://www.cisco.com/warp/public/459/13.html
- For more information on BGP solutions for scaling IBGP, refer to “BGP Case Studies Section 4” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#case4

Quiz: Configuring a Transit Autonomous System

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Identify the Cisco IOS commands required to configure IBGP neighbors
- Identify the Cisco IOS commands required to configure IBGP sessions between loopback interfaces
- Identify the Cisco IOS commands required to configure BGP synchronization to ensure successful IBGP operation of the transit AS
- Identify the Cisco IOS commands required to change the AD of BGP routes
- Identify the scalability limitations of IBGP-based backbones

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) When you are configuring the BGP neighbor session, what differentiates an EBGP neighbor from an IBGP neighbor?

A) The keyword **internal** at the end of the **neighbor** command.

B) IBGP neighbors will have the same AS number specified.

C) A description for the neighbor must be attached with the **neighbor description** command.

D) Directly connected neighbors will automatically form an EBGP session.

 Copyright © 2003, Cisco Systems, Inc.

Q2) What two steps are required to use a loopback interface for IBGP peering sessions? (Choose two.)

A) Ensure that the loopback interfaces are reachable through an IGP.

B) Ensure that the two neighbors must be directly attached.

C) Verify that each router has multiple physically redundant paths.

D) Configure a **neighbor** statement with the **update-source** command.

Q3) Why is it important to disable BGP synchronization in a transit backbone?

A) IGPs can support the routing requirements of full Internet routing, and hence synchronization is no longer necessary.

B) Because BGP redistribution into an IGP is no longer practical, enabling the synchronization feature is no longer applicable.

C) Synchronization requires all BGP transit routes to be explicitly mapped to an exit point, creating too much administrative overhead.

D) Synchronization requires BGP attributes to be properly mapped to IGP metrics in order for BGP routing across the transit backbone to function properly, creating too much overhead.

Q4) What are two negative ramifications of the full-mesh requirement imposed by IBGP? (Choose two.)

A) administratively difficult to apply an AS-wide routing policy

B) requires the use of next-hop-self for proper routing to external destinations

C) large number of TCP sessions

D) unnecessary duplication of routing traffic

Q5) What are two scalability tools that you can use to overcome the full-mesh requirement for IBGP sessions? (Choose two.)

A) confederations

B) floating static routes

C) route reflectors

D) disabling BGP synchronization

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring and Troubleshooting IBGP in Transit Autonomous Systems

Overview

Introduction of a transit backbone into a Border Gateway Protocol (BGP) network can create unique troubleshooting challenges. This lesson introduces Internal Border Gateway Protocol (IBGP) monitoring commands and troubleshooting techniques for solving the most common IBGP problems encountered in a transit backbone. Common problems with IBGP, as discussed in this lesson, occur when IBGP sessions do not reach the established state, routing information received via IBGP is never selected, and the best BGP route is never installed in the forwarding table.

Importance

Configuring a BGP network in a transit services configuration requires special care to ensure consistency of routing information throughout the autonomous system (AS). Understanding the tools and techniques to monitor and troubleshoot problems in the transit backbone is crucial to successfully implementing and maintaining the transit autonomous network.

Objectives

Upon completing this lesson, you will be able to:

- Identify the Cisco IOS® commands required to monitor IBGP operation
- Identify common IBGP configuration problems
- Troubleshoot IBGP session startup issues

- Troubleshoot IBGP route selection issues
- Troubleshoot IBGP synchronization issues

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

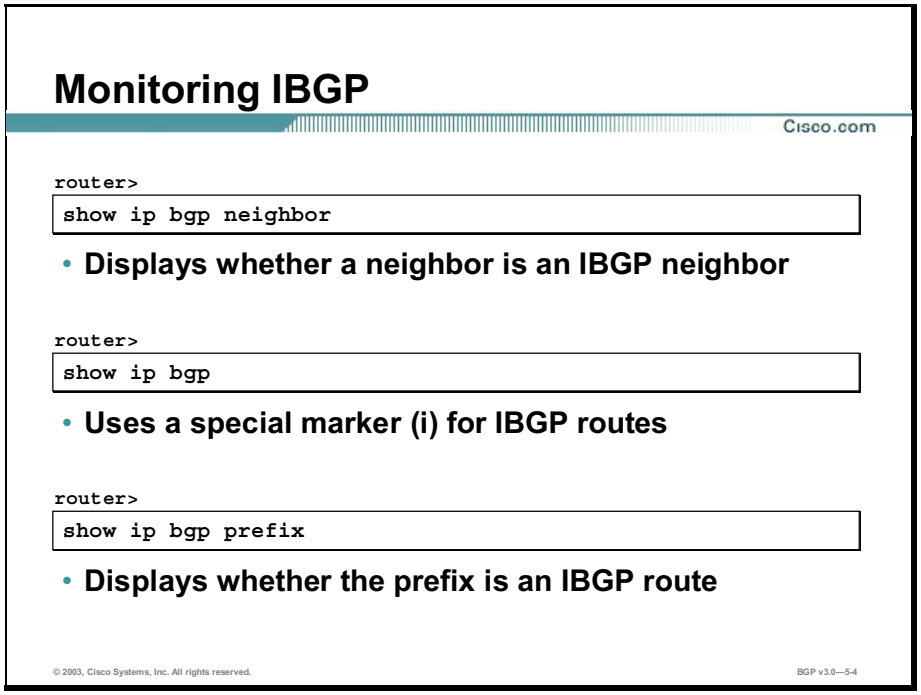
This lesson includes these topics:

- Overview
- Monitoring IBGP
- Common IBGP Problems
- Troubleshooting IBGP Session Startup Issues
- Troubleshooting IBGP Route Selection Issues
- Troubleshooting IBGP Synchronization Issues
- Summary
- Assessment (Lab): BGP Transit Autonomous Systems

 Copyright © 2003, Cisco Systems, Inc.

Monitoring IBGP

This topic lists the Cisco IOS commands required to monitor IBGP operation.


Monitoring IBGP
Cisco.com
router>
show ip bgp neighbor
• Displays whether a neighbor is an IBGP neighbor
router>
show ip bgp
• Uses a special marker (i) for IBGP routes
router>
show ip bgp prefix
• Displays whether the prefix is an IBGP route
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-4

show ip bgp neighbors

To display information about the TCP and BGP connections to neighbors, use the **show ip bgp neighbors** EXEC command.

show ip bgp neighbors [*ip-address*] [**received-routes** | **routes** | **advertised-routes** | {**paths** *regular-expression*} | **dampened-routes**]

Syntax Description

ip-address	(Optional) Address of the neighbor to display neighbor information about. If you omit this argument, all neighbors are displayed.
received-routes	(Optional) Displays all received routes (both accepted and rejected) from the specified neighbor.
routes	(Optional) Displays all routes that are received and accepted. The display output when using this keyword is a subset of the output from the **received-routes** keyword.
advertised-routes	(Optional) Displays all the routes that the router has advertised to the neighbor.
paths *regular-expression*	(Optional) Regular expression that the router uses to match the paths received.

dampened-routes	(Optional) Displays the dampened routes to the neighbor at the IP address specified

show ip bgp

To display entries in the BGP routing table, use the **show ip bgp** EXEC command.

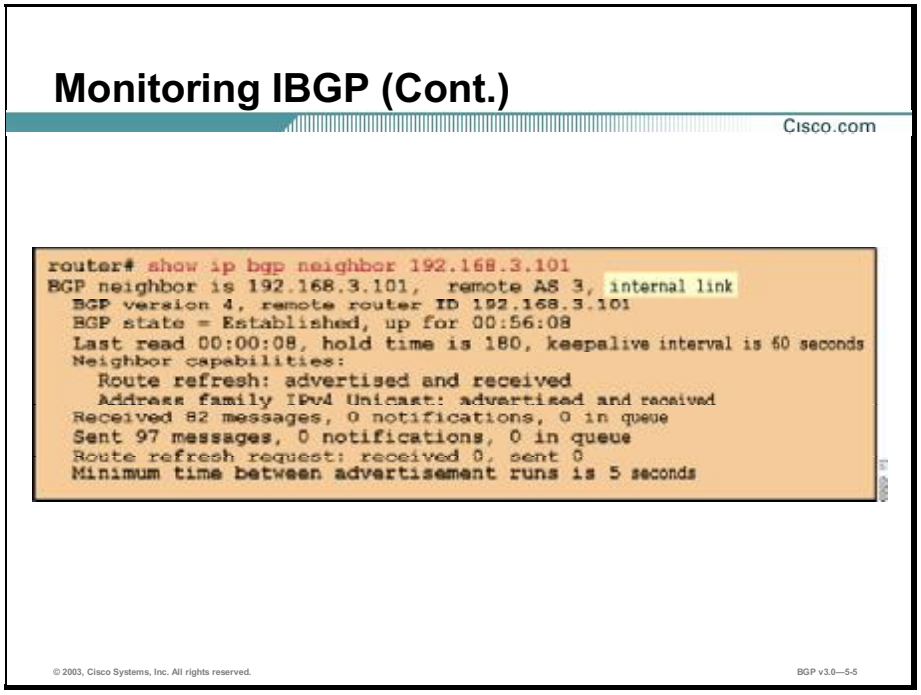
show ip bgp [*network*] [*network-mask*] [**longer-prefixes**]

Syntax Description

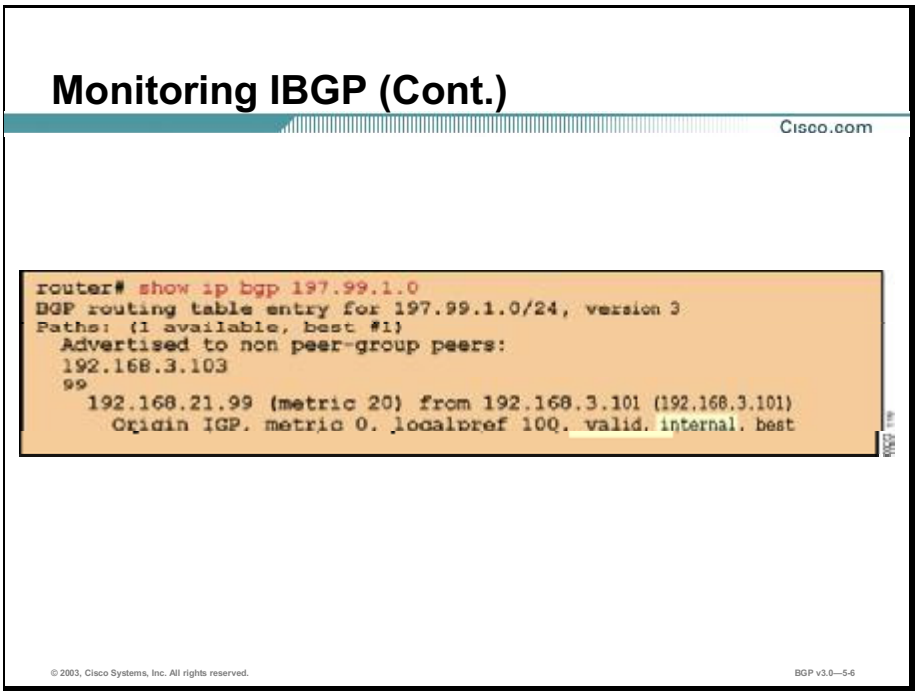
network	(Optional) Network number, entered to display a particular network in the BGP routing table
network-mask	(Optional) Displays all BGP routes matching the address-mask pair
longer-prefixes	(Optional) Displays the route and more specific routes

 Copyright © 2003, Cisco Systems, Inc.

Example


Monitoring IBGP (Cont.)
Cisco.com
router# show ip bgp neighbor 192.168.3.101
BGP neighbor is 192.168.3.101, remote AS 3, internal link
BGP version 4, remote router ID 192.168.3.101
BGP state = Established, up for 00:56:08
Last read 00:00:08, hold time is 180, keepalive interval is 60 seconds
Neighbor capabilities:
Route refresh: advertised and received
Address family IPv4 Unicast: advertised and received
Received 82 messages, 0 notifications, 0 in queue
Sent 97 messages, 0 notifications, 0 in queue
Route refresh request: received 0, sent 0
Minimum time between advertisement runs is 5 seconds
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-5

The **show ip bgp neighbors** command displays whether a router is running an IBGP or EBGP session with a BGP neighbor. The indication is given by the “internal link” phrase (marked yellow in the second line of the figure).


Monitoring IBGP (Cont.)
Cisco.com
router# show ip bgp 197.99.1.0
BGP routing table entry for 197.99.1.0/24, version 3
Paths: (1 available, best #1)
Advertised to non peer-group peers:
192.168.3.103
99
192.168.21.99 (metric 20) from 192.168.3.101 (192.168.3.101)
Origin IGP, metric 0, localpref 100, valid, internal, best
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-6

The **show ip bgp** *prefix* command displays whether a BGP route was received from an IBGP or EBGP neighbor. The indication is given by the keyword **internal** displayed in the last line of the printout (marked yellow in the last line of the figure).

Practice

Q1) Which Cisco IOS **show** command indicates that a BGP route is an IBGP route?

A) **show ip route**

B) **show ip route bgp**

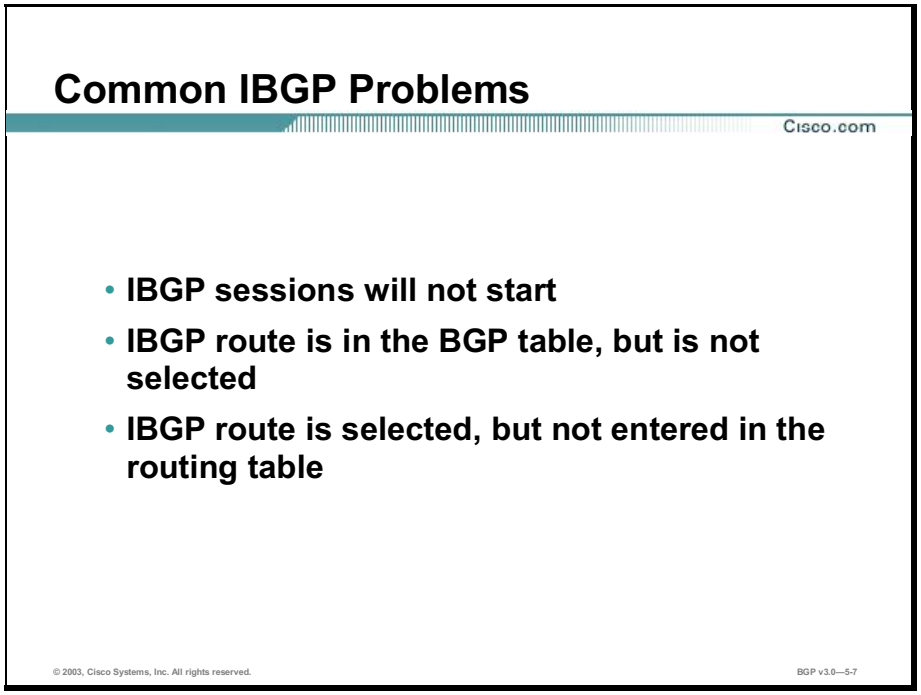
C) **show ip bgp**

D) **show ip bgp internal**

 Copyright © 2003, Cisco Systems, Inc.

Common IBGP Problems

This topic identifies configuration problems common to IBGP implementations.


Common IBGP Problems
Cisco.com
• IBGP sessions will not start
• IBGP route is in the BGP table, but is not selected
• IBGP route is selected, but not entered in the routing table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-7

Troubleshooting the BGP configuration of a transit AS can be cumbersome, because there are a number of common pitfalls that you might encounter. The next three topics give you troubleshooting advice to the most common problems:

- IBGP sessions do not reach the established state.
- Routing information received via IBGP is never selected.
- The best BGP route is never installed in the forwarding table.

Practice

Q1) What are the three common IBGP problems in transit backbones? (Choose three.)

A) IBGP sessions will not become established.

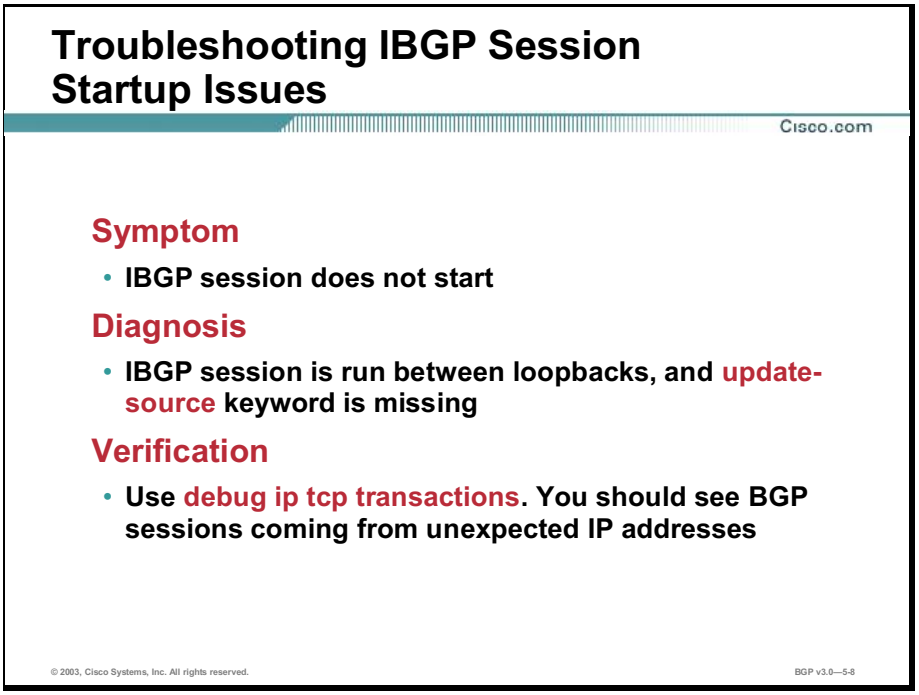
B) IBGP routes are never selected.

C) IBGP routes enter the “stuck in active state.”

D) BGP routes are not installed into the IP routing table.

Troubleshooting IBGP Session Startup Issues

This topic describes how to troubleshoot IBGP session startup issues.

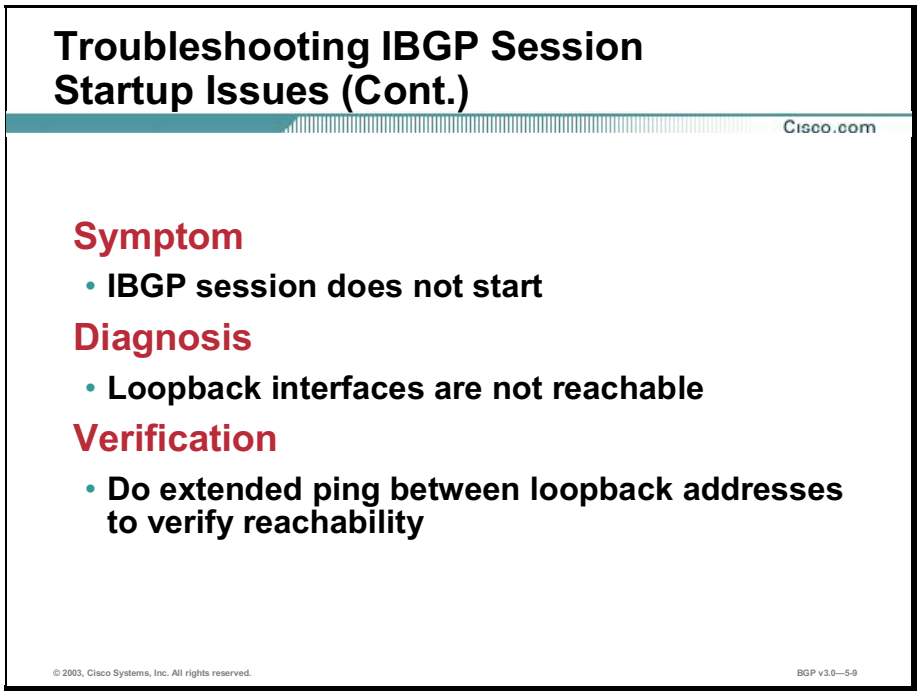

Troubleshooting IBGP Session Startup Issues
Cisco.com
Symptom
• IBGP session does not start
Diagnosis
• IBGP session is run between loopbacks, and update-source keyword is missing
Verification
• Use debug ip tcp transactions. You should see BGP sessions coming from unexpected IP addresses
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-8

A common mistake when you are configuring IBGP sessions is to forget the **neighbor update-source loopback 0** configuration command.

When configuring IBGP neighbors on the router, it is easy to remember to make a correct reference to the loopback interface of the remote router. But it is equally important to make sure that the correct source IP address of the outgoing TCP session is set. The peer router will not accept the session if the incoming source address does not match the peer router list of IBGP neighbors.

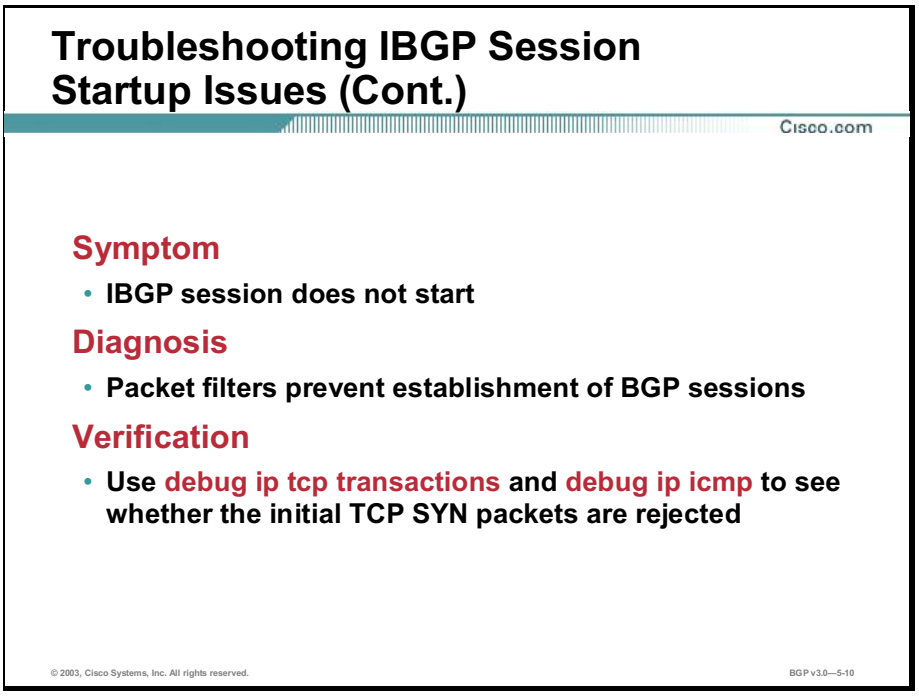
To verify that this situation is the problem, use the **debug ip tcp transactions** command. The output of the **debug ip tcp transactions** command should display TCP SYN packets coming from unexpected IP addresses on the receiving router and TCP sessions being reset with TCP RST packets on the sending (misconfigured) router.

 Copyright © 2003, Cisco Systems, Inc.


Troubleshooting IBGP Session Startup Issues (Cont.)
Cisco.com
Symptom
• IBGP session does not start
Diagnosis
• Loopback interfaces are not reachable
Verification
• Do extended ping between loopback addresses to verify reachability
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-9

An IBGP session between two routers can be established from the loopback interface of one router to the loopback interface of the other router only if the two routers can exchange IP packets using those addresses as source and destination. This exchange is possible only if the internal gateway protocol (IGP) carries the subnets assigned to each of the loopback interfaces.

When verifying the reachability with the **ping** command, make sure that the ping packets are sourced from the loopback interface. Use an extended **ping** and explicitly refer to the IP address of the loopback interface to ensure that packets are sourced from the loopback interface.


Troubleshooting IBGP Session Startup Issues (Cont.)
Cisco.com
Symptom
• IBGP session does not start
Diagnosis
• Packet filters prevent establishment of BGP sessions
Verification
• Use debug ip tcp transactions and debug ip icmp to see whether the initial TCP SYN packets are rejected
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-10

Packet filters can stop the BGP sessions. The path between the two BGP peer routers must be free from filters blocking the BGP traffic.

BGP runs on the well-known TCP port 179. Both routers will make connection attempts to that destination port. They will use a high-numbered TCP port as source. It is enough that one of the connection attempts succeed. But for better performance during recovery from network failure, both attempts should have the possibility to succeed. If both attempts do succeed, one of the connections will be brought down.

Practice

Q1) What are three common situations that prevent IBGP sessions from starting? (Choose three.)

A) The IBGP session has been configured to peer to a loopback interface, but **update-source** has not been configured on the neighbor.

B) An access control list filter is blocking access to TCP port 179.

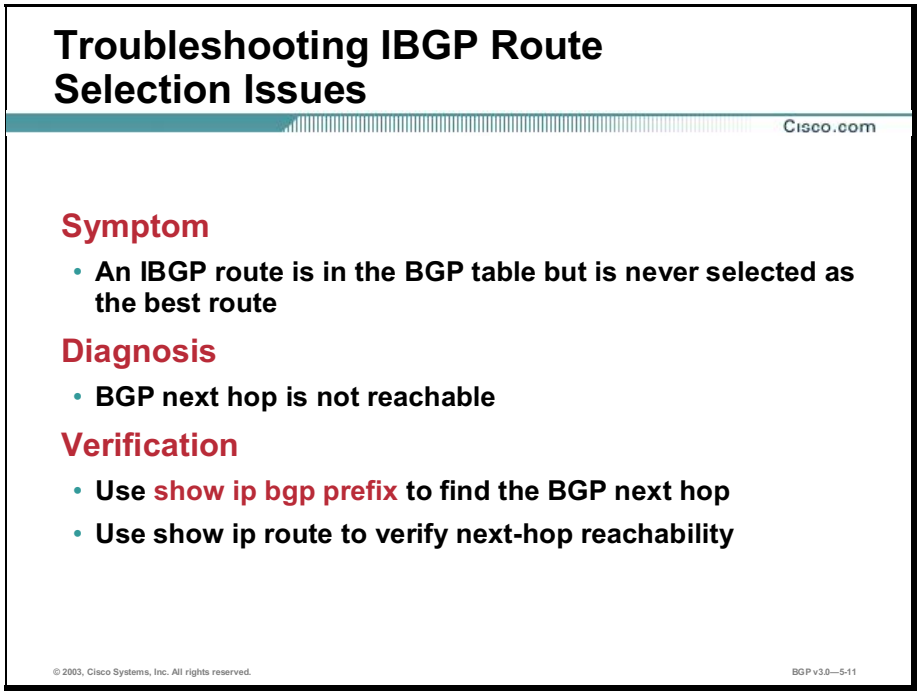
C) The IBGP session has been configured to peer to a loopback interface, but the loopback interface has not been administratively enabled with the **no shutdown** command.

D) The IBGP session has been configured to peer to a loopback interface, but the interfaces are not reachable via the IGP.

 Copyright © 2003, Cisco Systems, Inc.

Troubleshooting IBGP Route Selection Issues

This topic describes how to troubleshoot IBGP route selection issues.


Troubleshooting IBGP Route Selection Issues
Cisco.com
Symptom
• An IBGP route is in the BGP table but is never selected as the best route
Diagnosis
• BGP next hop is not reachable
Verification
• Use show ip bgp prefix to find the BGP next hop
• Use show ip route to verify next-hop reachability
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-11

A BGP update can be used by the router to reach network destinations only if the next-hop address specified in the BGP update is reachable. A BGP update, which refers to a next hop that is currently not reachable according to the forwarding table, will be saved in the BGP table, but it cannot be installed by the router into its forwarding table. If the next-hop address later becomes reachable, the BGP route will become a candidate route that could be used by that router for packet forwarding to that destination.

To verify the next-hop reachability, check the BGP route in the BGP table using the **show ip bgp** *prefix* command. The next hop is referred to as "inaccessible," if it is not currently reachable according to the forwarding table.

A common mistake is to forget to let the IGP announce the reachability of subnets that physically connect the local AS with a neighboring AS. These subnets are used by the router to establish the EBGP session, and the next hop received in an incoming BGP update will be the far end of the EBGP session. If all routers in the local AS do not have a path to that subnet, the next-hop address will be inaccessible.

Prevent this problem by including the subnet linking the transit AS to neighboring autonomous systems in the IGP using either the **redistribute connected** or the **network and passive-interface** configuration commands.

Practice

Q1) What would prevent IBGP routes from being selected as the best route in the BGP table?

A) Failure to disable BGP synchronization.

B) Failure to disable BGP split horizon.

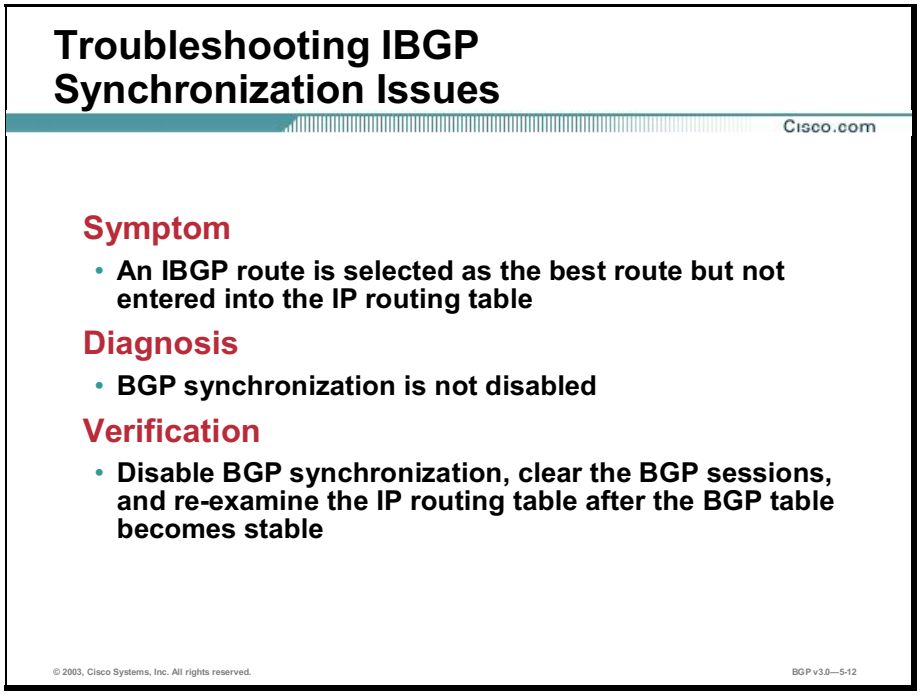
C) The IGP has no route to the BGP next hop.

D) A default route has not been injected into the IGP.

 Copyright © 2003, Cisco Systems, Inc.

Troubleshooting IBGP Synchronization Issues

This topic describes how to troubleshoot IBGP synchronization issues.


Troubleshooting IBGP Synchronization Issues
Cisco.com
Symptom
• An IBGP route is selected as the best route but not entered into the IP routing table
Diagnosis
• BGP synchronization is not disabled
Verification
• Disable BGP synchronization, clear the BGP sessions, and re-examine the IP routing table after the BGP table becomes stable
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-12

In old BGP designs, redistribution between BGP and IGP was common practice, and these protocols had to be synchronized to ensure proper packet forwarding. In modern designs, redistribution is no longer used and the synchronization has to be turned off. However, the default value is to have synchronization enabled.

Routers with BGP synchronization enabled will not install IBGP routes in the forwarding table nor propagate them to other EBGP neighbors.

Fix this problem by configuring **no synchronization** in the router BGP configuration.

Practice

Q1) What common issue could prevent IBGP best routes from being inserted into the IP routing table?

A) Failure to disable BGP synchronization.

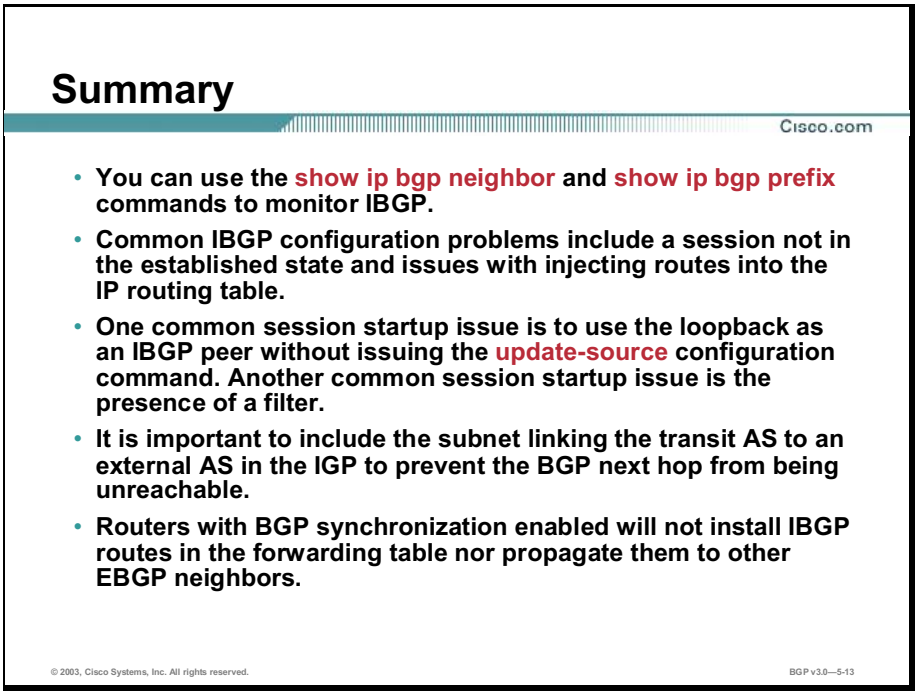
B) Failure to disable BGP split horizon.

C) The IGP has no route to the BGP next hop.

D) A default route has not been injected into the IGP.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• You can use the show ip bgp neighbor and show ip bgp prefix commands to monitor IBGP.
• Common IBGP configuration problems include a session not in the established state and issues with injecting routes into the IP routing table.
• One common session startup issue is to use the loopback as an IBGP peer without issuing the update-source configuration command. Another common session startup issue is the presence of a filter.
• It is important to include the subnet linking the transit AS to an external AS in the IGP to prevent the BGP next hop from being unreachable.
• Routers with BGP synchronization enabled will not install IBGP routes in the forwarding table nor propagate them to other EBGP neighbors.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—5-13

Next Steps

After completing this lesson, go to:

- Scaling Service Provider Networks module

References

For additional information, refer to these resources:

- For more information on troubleshooting IBGP, refer to “Troubleshooting IP Connectivity and Routing Problems” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/itg_v1/tr1907.htm#xtocid27

- For more information on configuring and monitoring IBGP, refer to “Configuring BGP” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2/1cfbgp.htm

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: BGP Transit Autonomous Systems

Complete the laboratory exercise to practice what you have learned in this module.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Copyright © 2003, Cisco Systems, Inc.

Exercise Objective

In this exercise, you will enable the provider network to behave as a transit AS, given a typical service provider network with multiple BGP connections to other autonomous systems.

After completing this exercise, you will be able to:

- Plan the migration of an existing backbone toward a fully meshed IBGP backbone designed for transit traffic
- Configure IBGP sessions between loopback interfaces
- Configure BGP synchronization to ensure successful IBGP operation of the transit AS
- Monitor IBGP operation

Command List

The commands used in this exercise are described in the table here.

Table 1: Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor *ip-address* **remote-as** *as-number*	Establish an IBGP session by using your workgroup number as the AS number.
neighbor *ip-address* **update-source** *interface*	Use IP address of the specified interface' as the source address for the BGP session.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.
no synchronization	Disable synchronization of IGP and BGP routes.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- With the rapid growth of the Internet, you decide to become an Internet service provider (ISP), and you already have your first customer. Unfortunately, the customer is willing to pay you only for connectivity toward your own network and toward AS 37.
- In this exercise, you will transform your network into a transit AS running BGP on all core routers.
- Start a BGP process on all routers in your workgroup. Configure a full mesh of IBGP sessions between all routers in your AS. Establish these BGP sessions between Loopback0 interfaces.

 Copyright © 2003, Cisco Systems, Inc.

- Propagate only your own networks and networks originating in AS 37 to router “Client.”
- Make sure that you accept only the networks originating in AS 99 from router “Client.”
- AS 99 should receive only prefixes originating in your AS “x” and AS 37. You should accept only prefixes originating in AS 99 from router “Client.”
- Figure 1 displays the required BGP connectivity within your AS as well as the BGP sessions with your customer and your upstream ISPs.

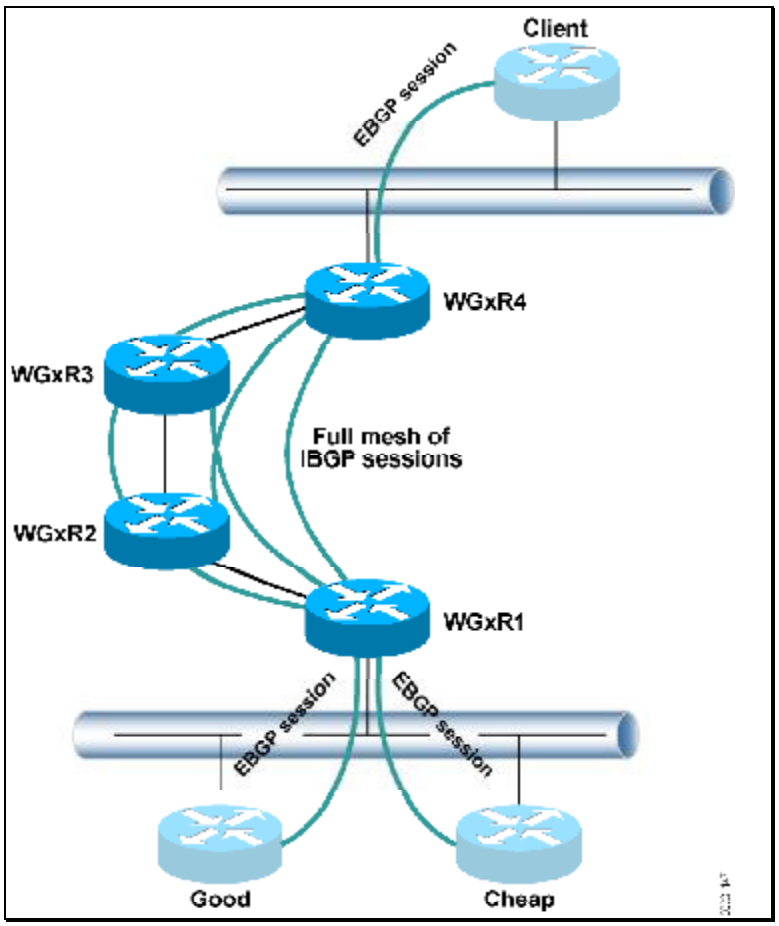


Figure 1: Creating a full mesh of IBGP sessions

Task 1: Configure the BGP Transit Autonomous System

In this task, you will configure your network backbone as a fully meshed IBGP backbone acting as a transit AS.

Exercise Procedure

Complete these steps:

Step 1 To clean up your BGP configuration, remove the BGP process and the default route from IGP on WGxR1.

Step 2 Remove the BGP process from WGxR2.

Step 3 Start the BGP process on all routers in your workgroup.

Step 4 Advertise your prefixes (197.x.0.0/16 and 192.168.x.0/24) in BGP on routers WGxR1 and WGxR4.

Step 5 Re-establish neighbor relationships to routers "Good" and "Cheap" without any filters on router WGxR1. Use the parameters from the following table:

Router	AS number	IP address
Good	20	192.168.20.20
Cheap	22	192.168.20.22

Step 6 Establish a BGP session with router "Client" on router WGxR4. Use the parameters from the following table:

Parameter	Value
Client IP address	192.168.21.99
Client AS number	99

Step 7 Configure all routers in your workgroup as IBGP neighbors (IBGP full mesh). Use loopback interfaces to establish these IBGP sessions.

Exercise Verification

You have completed this exercise when you attain these results:

- Check BGP on all core routers and the router "Client" and ensure that they have established the correct sessions with their peers.

```
WG1R1#sh ip bgp summary
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.20.20   4    20    1189    1200       52    0    0 13:34:45
23
192.168.20.22   4    22    1195    1183       52    0    0 13:34:46
23
197.1.2.1       4     1    1174    1196       52    0    0 13:34:58
0
```

 Copyright © 2003, Cisco Systems, Inc.

```
197.1.4.1       4     1    1170    1188        52    0    0 13:35:17
0
197.1.6.1       4     1    1171    1183        52    0    0 13:34:58
13
```

```
WG1R2#sh ip bgp summary
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
197.1.1.1       4     1    1200    1179      125    0    0 13:38:09
24
197.1.4.1       4     1    1173    1173      125    0    0 13:38:31
0
197.1.6.1       4     1    1176    1170      125    0    0 13:37:58
13
```

```
WG1R3#sh ip bgp sum
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
197.1.1.1       4     1    1193    1175       78    0    0 13:40:33
24
197.1.2.1       4     1    1175    1175       78    0    0 13:40:37
0
197.1.6.1       4     1    1183    1175       78    0    0 13:40:04
13
```

```
WG1R4#sh ip bgp sum
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.21.99   4    99    1191    1192       14    0    0 13:40:38
11
197.1.1.1       4     1    1190    1178       14    0    0 13:41:04
24
197.1.2.1       4     1    1173    1179       14    0    0 13:40:54
0
197.1.4.1       4     1    1175    1183       14    0    0 13:40:54
0
```

```
Client#sh ip bgp sum
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.21.X    4     1    1147    1146       18    0    0 13:42:01
2
```

- Check the BGP table on router “Client” and verify that it is correctly receiving BGP routes.

```
Client#sh ip bgp
BGP table version is 119, local router ID is 197.99.111.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.0.0.0         192.168.21.1                           0 1 20 i
*> 99.0.0.0         0.0.0.0                  0         32768 i
*> 128.20.0.0       192.168.21.1                           0 1 22 i
*> 128.20.12.0/24   192.168.21.1                           0 1 20 i
*> 128.22.0.0       192.168.21.1                           0 1 22 i
*> 128.22.12.0/24   192.168.21.1                           0 1 22 i
*> 128.26.0.0       192.168.21.1                           0 1 22 26
i
*> 128.37.0.0       192.168.21.1                           0 1 20 42
37 i
*> 128.42.0.0       192.168.21.1                           0 1 20 42
i
*> 128.51.0.0       192.168.21.1                           0 1 22 26
51 i
*> 128.213.0.0      192.168.21.1                           0 1 20 213
i
*> 128.214.0.0      192.168.21.1                           0 1 22 214
i
*> 192.20.11.0      192.168.21.1                           0 1 22 i
*> 192.20.12.0/30   192.168.21.1                           0 1 20 i
*> 192.22.11.0      192.168.21.1                           0 1 22 i
*> 192.22.12.0/30   192.168.21.1                           0 1 22 i
*> 192.26.11.0      192.168.21.1                           0 1 22 26
i
*> 192.37.11.0      192.168.21.1                           0 1 20 42
37 i
*> 192.42.11.0      192.168.21.1                           0 1 20 42
i
*> 192.51.11.0      192.168.21.1                           0 1 22 26
51 i
*> 192.168.1.0      192.168.21.1             0             0 1 i
*> 192.213.11.0     192.168.21.1                           0 1 20 213
i
*> 192.214.11.0     192.168.21.1                           0 1 22 214
i
*> 197.1.0.0/16     192.168.21.1             0             0 1 i
*> 197.99.1.0       0.0.0.0                  0         32768 i
*> 197.99.11.0      0.0.0.0                  0         32768 i
*> 197.99.12.0      0.0.0.0                  0         32768 i
*> 197.99.13.0      0.0.0.0                  0         32768 i
*> 197.99.20.0      0.0.0.0                  0         32768 i
*> 197.99.22.0      0.0.0.0                  0         32768 i
*> 197.99.80.0      0.0.0.0                  0         32768 i
*> 197.99.111.1/32  0.0.0.0                  0         32768 i
```

 Copyright © 2003, Cisco Systems, Inc.

```
*> 197.99.120.0     0.0.0.0                  0         32768 i
*> 197.99.128.0/20  0.0.0.0                  0         32768 i
*> 200.20.0.0/16    192.168.21.1                           0 1 20 i
*> 200.22.0.0/16    192.168.21.1                           0 1 22 i
```

- Use traceroute from router “Client” to the loopback interface on router WGxR1 (197.x.1.1). You should see a path similar to the one below:

```
Client# traceroute 197.1.1.1
Type escape sequence to abort.
Tracing the route to 197.1.1.1

  1 192.168.21.1 4 msec 4 msec 4 msec
  2 192.168.1.9 [AS 1] 20 msec 16 msec 16 msec
  3 192.168.1.5 [AS 1] 32 msec 32 msec 28 msec
  4 192.168.1.1 [AS 1] 44 msec *  44 msec
```

- Use traceroute from router WGxR1 to the loopback interface on router “Client” (197.99.1.1). You should see a path similar to the one below:

```
WG1R1# traceroute 197.99.1.1
Type escape sequence to abort.
Tracing the route to 197.99.1.1

  1 192.168.1.2 16 msec 16 msec 17 msec
  2 192.168.1.6 32 msec 32 msec 28 msec
  3 192.168.1.10 44 msec 40 msec 40 msec
  4 192.168.21.99 48 msec *  44 msec
```

Answer these questions:

Q1) Check the BGP table on router “Client.” How many prefixes coming from your AS are in that BGP table? _____

Q2) Is there any other way of discovering how many prefixes you have advertised to the router “Client”? _____

Task 2: Configure Filters in the BGP Transit Autonomous System

As the last steps in this exercise, you need to establish route filters toward your customer on WGxR4.

Exercise Procedure

Complete these steps:

Step 1 Create an AS-path filter to permit your own networks and networks originating in AS 37.

Step 2 Create an AS-path filter to permit networks originating in AS 99.

Step 3 Apply the AS-path filters to your customer.

Exercise Verification

You have completed this exercise when you attain these results:

- Check to see if router “Client” is receiving only your networks and those originating in AS 37.

```
Client#sh ip bgp
BGP table version is 206, local router ID is 197.99.111.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 99.0.0.0         0.0.0.0                  0         32768 i
*> 128.37.0.0       192.168.21.1                           0 1 20 42
37 i
*> 192.37.11.0      192.168.21.1                           0 1 20 42
37 i
*> 192.168.1.0      192.168.21.1             0             0 1 i
*> 197.1.0.0/16     192.168.21.1             0             0 1 i
*> 197.99.1.0       0.0.0.0                  0         32768 i
*> 197.99.11.0      0.0.0.0                  0         32768 i
*> 197.99.12.0      0.0.0.0                  0         32768 i
*> 197.99.13.0      0.0.0.0                  0         32768 i
*> 197.99.20.0      0.0.0.0                  0         32768 i
*> 197.99.22.0      0.0.0.0                  0         32768 i
*> 197.99.80.0      0.0.0.0                  0         32768 i
*> 197.99.111.1/32  0.0.0.0                  0         32768 i
*> 197.99.120.0     0.0.0.0                  0         32768 i
*> 197.99.128.0/20  0.0.0.0                  0         32768 I
```

 Copyright © 2003, Cisco Systems, Inc.

- Use the **show ip route ospf** command to make sure that your IGP carries only your internal networks.

```
WG1R4#sh ip route ospf
     197.1.8.0/32 is subnetted, 1 subnets
O       197.1.8.1 [110/193] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.1.0/32 is subnetted, 1 subnets
O       197.1.1.1 [110/193] via 192.168.1.9, 1d02h, Serial0/0.1
O    192.168.20.0/24 [110/202] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.3.0/32 is subnetted, 1 subnets
O       197.1.3.1 [110/129] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.2.0/32 is subnetted, 1 subnets
O       197.1.2.1 [110/129] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.5.0/32 is subnetted, 1 subnets
O       197.1.5.1 [110/65] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.4.0/32 is subnetted, 1 subnets
O       197.1.4.1 [110/65] via 192.168.1.9, 1d02h, Serial0/0.1
     192.168.1.0/24 is variably subnetted, 4 subnets, 2 masks
O       192.168.1.0/30 [110/192] via 192.168.1.9, 1d02h, Serial0/0.1
O       192.168.1.4/30 [110/128] via 192.168.1.9, 1d02h, Serial0/0.1
```

Answer these questions:

Q1) Why did you have to disable synchronization?

Q2) Why did you have to establish a full mesh of IBGP sessions?

 Copyright © 2003, Cisco Systems, Inc.

6

Scaling Service Provider Networks

Overview

Common to every service provider network are concerns over network scalability, specifically when the network supports both internal applications and customer connectivity. When interior and exterior routing protocols are deployed in large, complex service provider networks, protocol interaction, interior gateway protocol (IGP) choice, and addressing are common concerns. Proper scaling of Internal Border Gateway Protocol (IBGP) within the service provider network is also an area of concern.

In standard Border Gateway Protocol (BGP) implementations, all BGP routers within an autonomous system (AS) must be fully meshed so that all external routing information can be distributed among the other routers residing within the AS. Therefore, within an AS, all routers must establish TCP sessions with all other BGP routers. As the AS grows, scalability challenges arise because of an ever-increasing number of TCP sessions and demands for router CPU and memory resources.

This module discusses network scalability concerns that are common to large, complex service provider networks. The module also discusses BGP route reflectors and confederations as scalability mechanisms that allow network designers to steer away from BGP full-mesh requirements and improve network scalability by reducing the number of TCP sessions required within an AS. Also discussed in this module are the Cisco IOS® commands needed to configure and monitor BGP route reflectors and confederations.

Upon completing this module, you will be able to:

- Describe common routing scalability concerns in service provider networks
- Describe the function and operation of route reflectors in a BGP environment
- Describe the concept of hierarchical route reflectors and their requirements, based upon established route reflector design rules

- Configure and verify proper operation of route reflectors to modify IBGP split-horizon rules, given an existing IBGP network
- Describe the function and operation of confederations in a BGP environment
- Configure and verify proper operation of confederations to modify IBGP AS-path processing, given an existing IBGP network

Outline

The module contains these lessons:

- Scaling IGP and BGP in Service Provider Networks
- Introduction to Route Reflectors
- Network Design with Route Reflectors
- Configuring and Monitoring Route Reflectors
- Introduction to Confederations
- Configuring and Monitoring Confederations

 Copyright © 2003, Cisco Systems, Inc.

Scaling IGP and BGP in Service Provider Networks

Overview

Common to every service provider network are concerns about network scalability. The interactions between interior gateway protocols (IGPs) and the Border Gateway Protocol (BGP), specifically when supporting internal routing, customer connectivity, and transit traffic (and the administrative policies that match), can be quite complex. Furthermore, the large number of prefixes required to support full Internet routing requires network administrators to fully characterize IGP and BGP interactions for internal networks and customers alike.

This lesson discusses network scalability concerns common to large, complex service provider networks. Included in this lesson is a description of a typical service provider network, and discussion of the propagation of internal and customer routing information, scaling considerations for IGPs and BGP, and scaling of IP addressing in service provider networks.

Importance

Properly scaling IP addressing, IGPs, and BGP is a common area of concern to all service providers and can be the difference between a successful and a problematic BGP implementation. Because service provider networks are complex and must meet the administrative policy and routing demands of the internal network, different customers, and other providers, proper scaling is crucial to the success of the network.

Objectives

Upon completing this lesson, you will be able to:

- Describe the basic structure of service provider networks
- Describe the propagation of internal and customer routes in service provider networks
- Describe proper scaling of IGPs and BGP in service provider networks
- Describe IP addressing scaling issues in service provider networks

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- BGP Transit Autonomous Systems module

Outline

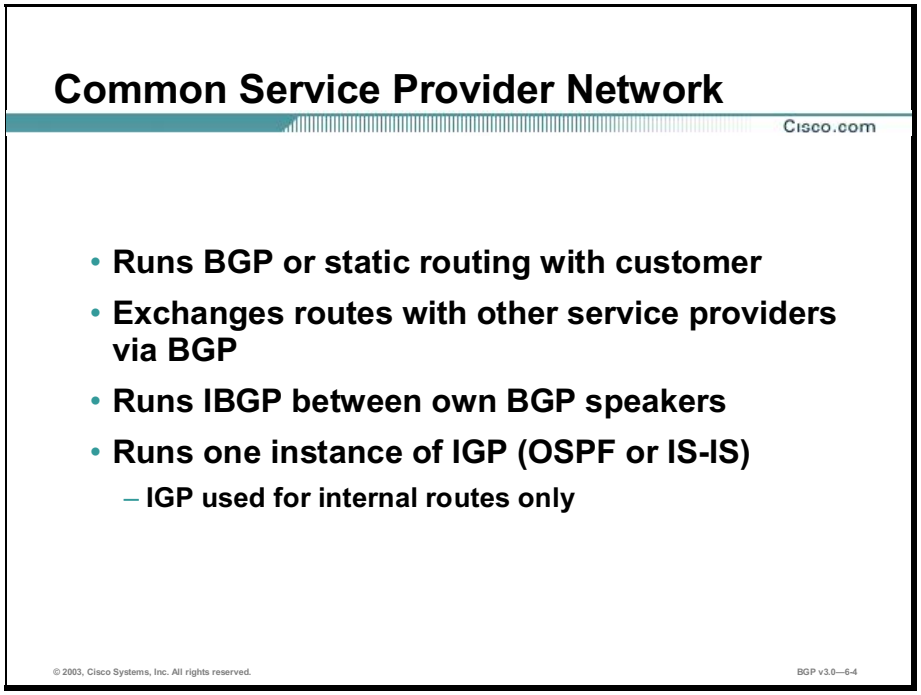
This lesson includes these topics:

- Overview
- Common Service Provider Network
- Route Propagation in Service Provider Networks
- Scaling Service Provider Routing Protocols
- Scaling Service Provider Addressing
- Summary
- Assessment (Quiz): Scaling IGP and BGP in Service Provider Networks

 Copyright © 2003, Cisco Systems, Inc.

Common Service Provider Network

This topic describes the basic structure of service provider networks.


Common Service Provider Network
Cisco.com
• Runs BGP or static routing with customer
• Exchanges routes with other service providers via BGP
• Runs IBGP between own BGP speakers
• Runs one instance of IGP (OSPF or IS-IS)
– IGP used for internal routes only
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-4

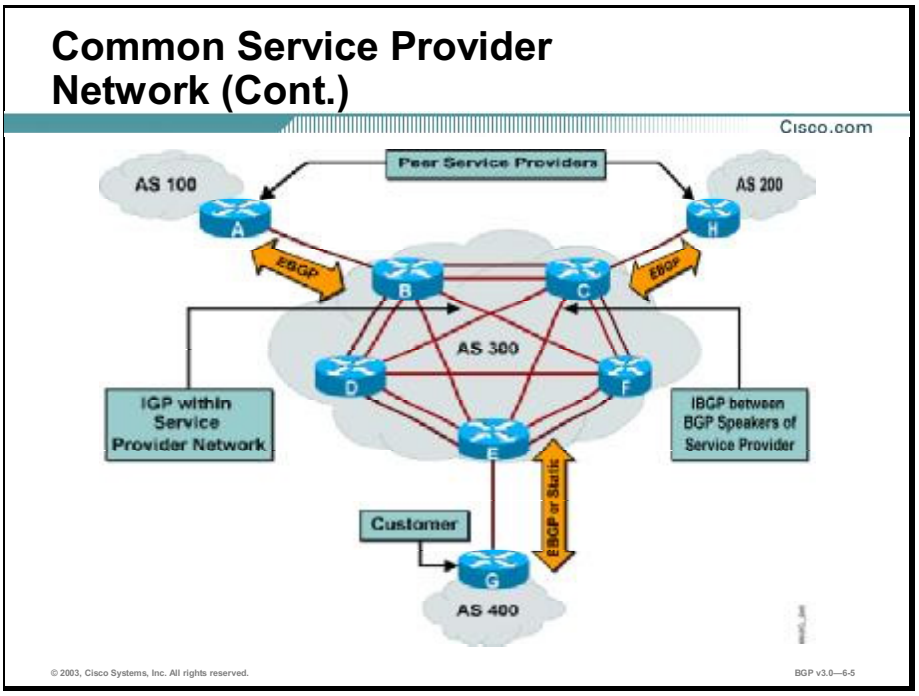
The common service provider network runs External Border Gateway Protocol (EBGP) or static routing with customers. EBGP is always used as the routing protocol between different service providers.

Internal Border Gateway Protocol (IBGP) is required in the provider network because all EBGP speaking routers in an autonomous system (AS) must exchange external routes via IBGP. Also non-EBGP speakers are required to take part in the IBGP exchange if they are in a transit path and forward packets based on destination IP addresses.

The service provider network also runs an IGP. The protocols of choice are Open Shortest Path First (OSPF) and Intermediate System-to-Intermediate System (IS-IS). The IGP is used for two purposes:

- Provides IP connectivity between all IBGP speakers so that TCP sessions for IBGP can be established between BGP speaking routers
- Provides optimal routing to the BGP next-hop address

A single IGP should be used within the entire AS. This setup facilitates effective packet forwarding from the ingress router to egress routers. The IGP is configured to carry internal routes only, including internal links and loopback addresses of the routers. However, for performance and scalability reasons, no customer routes or external routes should be injected into the IGP.



The typical service provider network consists of a network core that connects various edge devices. Some of the edge devices connect customers; others connect to other service providers.

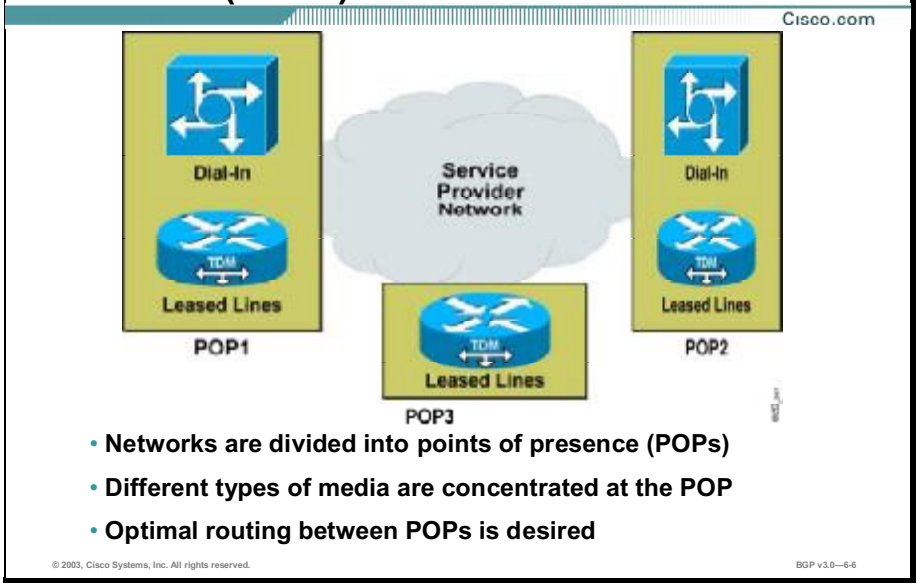
The edge devices connecting to other service providers use EBGP to exchange routing information. The edge devices connecting customers use either static routing or EBGP.

Unless Multiprotocol Label Switching (MPLS) is configured on the service provider backbone, routers in a transit path are also required to have full routing information. Therefore, these routers take part in the IBGP routing exchange.

An IGP is also required within the service provider network. The IGP is used to carry internal routes, including the loopback interface addresses of IBGP-speaking routers. The IGP provides reachability information to establish IBGP sessions and to perform the recursive routing lookup for the BGP next hop.

 Copyright © 2003, Cisco Systems, Inc.

Common Service Provider Network (Cont.)

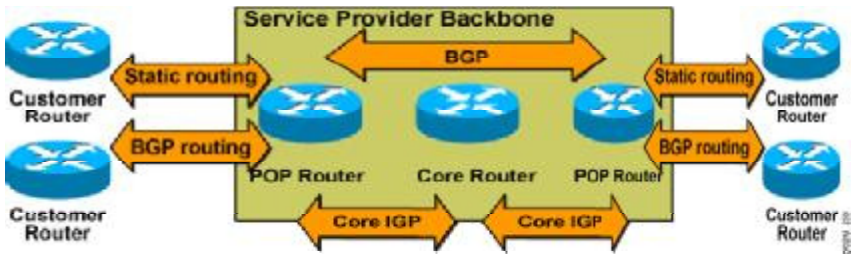


Service provider devices that connect access links to customers are physically located in groups, called points of presence (POPs). In general, the POP is a group of routers where access links are terminated. The edge routers peering with other service providers can in this sense be considered a POP.

Service providers use different types of access links with different types of customers and usually mix access links in the same POP. Some customers use leased lines, others use xDSL, and still others use dial-in access or any other access that the provider can support.

POP routers connect to the network core using a layer of concentration routers at the POP. The network core forwards packets between POPs, different customer access points, or peering points with other service providers. Optimal routing between POPs is a desirable feature.

Common Service Provider Network (Cont.)



- **POP routers use BGP or static routing with customer routers**
- **Provider core IGP is a single instance of IS-IS or OSPF**
- **The core IGP is used only within the service provider backbone**

© 2003, Cisco Systems, Inc. All rights reserved.

Customer access lines terminate in the POP edge routers. In many cases, the POP edge routers use static routing to customer networks. The POP edge routers advertise static routes to the rest of the service provider network and to other autonomous systems using BGP.

Service providers use BGP routing with the customer when redundancy requires the use of a routing protocol.

The service provider backbone typically uses a single instance of either IS-IS or OSPF as its IGP. The IGP is used within the provider backbone only. The provider backbone exchanges no IGP routing information with customer routers or with routers in other autonomous systems.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) How is routing information exchanged between the service provider and other autonomous systems?

A) IBGP

B) EBGP

C) static routes

D) with an IGP

Q2) In what two ways is routing information typically exchanged between a service provider and its customers? (Choose two.)

A) IBGP

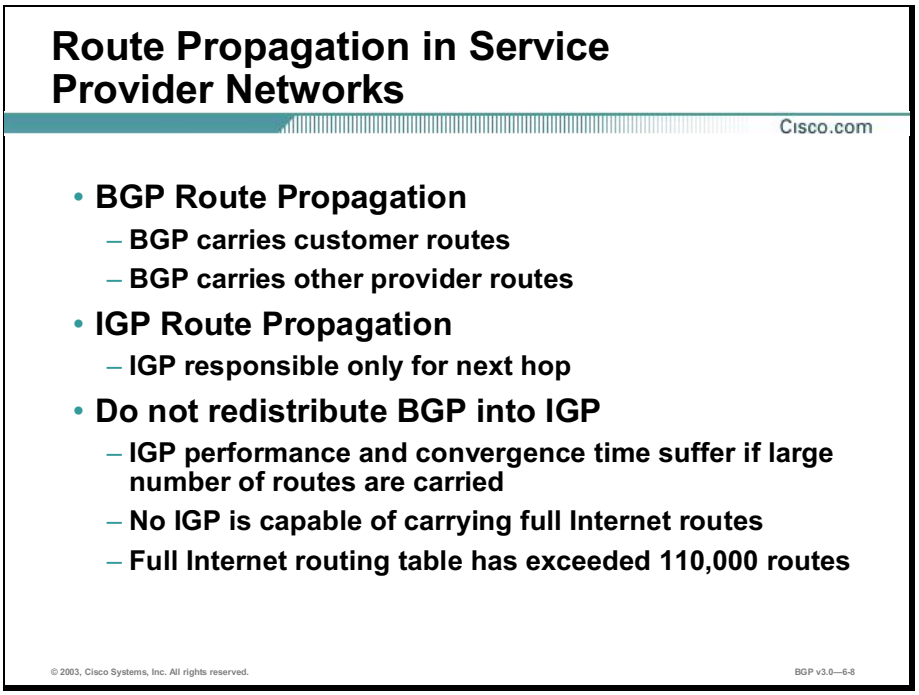
B) EBGP

C) static routes

D) with an IGP

Route Propagation in Service Provider Networks

This topic describes the propagation of internal and customer routes in service provider networks.


Route Propagation in Service Provider Networks
Cisco.com
• BGP Route Propagation
– BGP carries customer routes
– BGP carries other provider routes
• IGP Route Propagation
– IGP responsible only for next hop
• Do not redistribute BGP into IGP
– IGP performance and convergence time suffer if large number of routes are carried
– No IGP is capable of carrying full Internet routes
– Full Internet routing table has exceeded 110,000 routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-8

It is important to avoid sending any unnecessary routing information in the IGP. The IGP performs best if it carries as few routes as possible. Optimally, the IGP should contain only information about BGP next hops and routes internal to the service provider network, enabling the establishment of IBGP sessions.

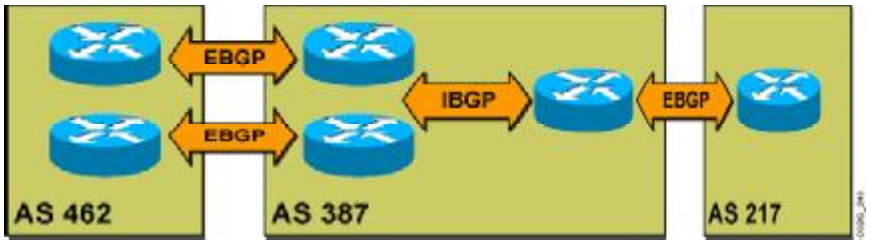
All other routing information should be carried in BGP, which is designed to scale for large volumes of routing information. Customer routes and the routes from other service providers should be carried in BGP. These routes should not be propagated from BGP into the provider IGP.

IGP performance and convergence time suffer if IGP carries a larger number of routes. The design goal should be to minimize the volume of routing information carried by the IGP. Naturally, the number of route flaps is also reduced as the number of routes is reduced.

BGP scales to a much larger volume of routing information due to the inherent qualities of the design of BGP. Potentially, the BGP routers of the service provider could receive the full Internet routing table, which has exceeded 110,000 routes. You should therefore never redistribute the routing information received by BGP into the IGP, because no IGP is capable of carrying several tens of thousands of routes.

 Copyright © 2003, Cisco Systems, Inc.

Routing Information Exchange with Other Service Providers



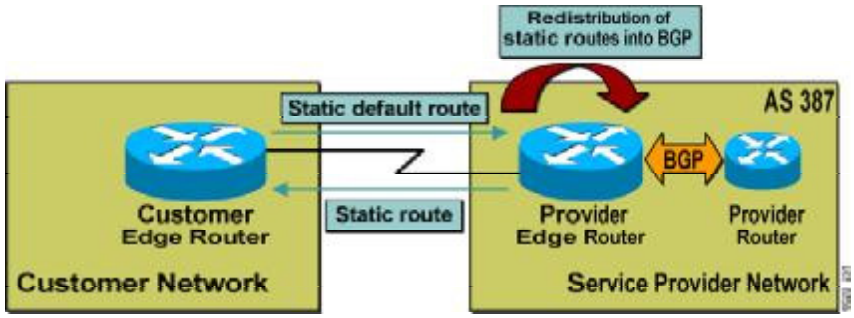
- **BGP is used to exchange routing information between Internet service providers**

© 2003, Cisco Systems, Inc. All rights reserved.

Provider edge routers use BGP to exchange routing information with other service provider networks for redundancy and scalability reasons.

Static routing with other service providers is generally not a viable solution due to the dynamic routing requirements of the service provider environment. Routing information is received at provider edge routers using EBGP and then propagated using IBGP to the rest of the service provider network. At another edge router, the routing information is further propagated to a different service provider using EBGP with other autonomous systems.

Routing Information Exchange with Customers



- **The provider edge router redistributes static customer routes into BGP**
- **BGP carries customer routes**

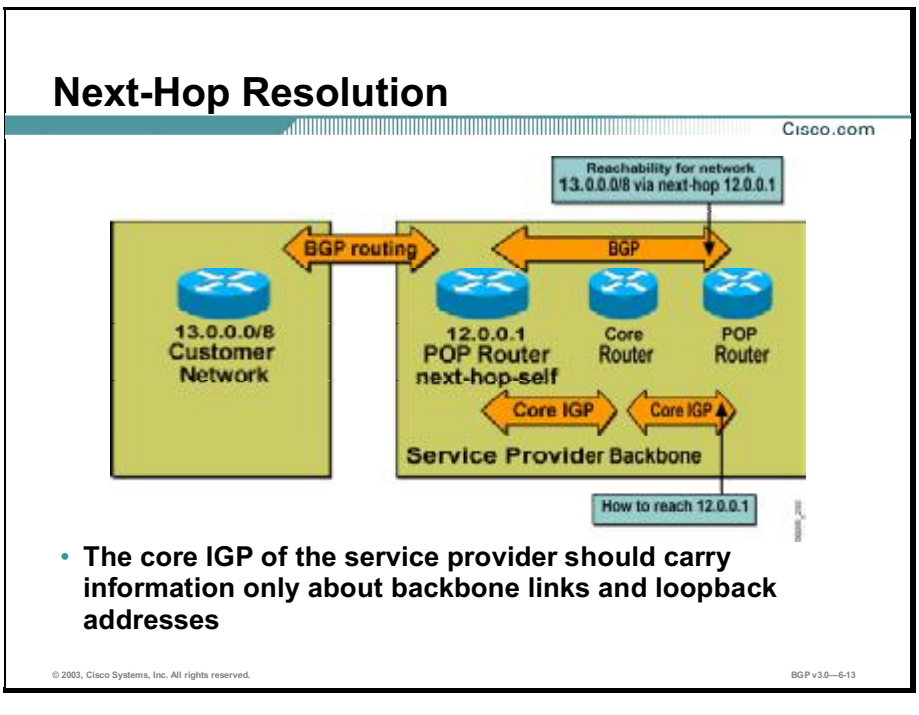
© 2003, Cisco Systems, Inc. All rights reserved.

The provider edge router typically uses static routing to reach customer networks. In this case, the customer typically configures a static default route pointing to the edge router of the service provider.

The provider edge router redistributes customer static routes into BGP. The service provider network then uses BGP to propagate the information to the rest of the service provider network using IBGP. The service provider also advertises customer routing information to other autonomous systems using BGP.

 Copyright © 2003, Cisco Systems, Inc.

Next-Hop Resolution



The IGP used in the service provider core should carry information only about backbone links and loopback addresses. The service provider should use BGP to carry all other information.

Use BGP **next-hop-self** when BGP routing is exchanged with the customer or other service providers. Using **next-hop-self** results in the BGP next hop being set to the loopback address of the service provider edge router and not to the access link address of the customer. The IGP can then be relieved from the burden of carrying information about the access link. The benefit of not carrying customer link information is that a flapping access link will not disturb the service provider IGP.

Practice

Q1) What would be the impact of redistributing BGP into the provider IGP?

A) The transit backbone would stop functioning because all IGPs have smaller administrative distances than IBGP.

B) Convergence times would improve because IGPs have faster convergence and smaller advertisement intervals than BGP.

C) Unbounded, the large number of routes carried in BGP would exceed the scalability of any IGP.

D) There is little to no impact when redistributing BGP routes into the provider IGP.

Q2) What benefit is gained if links to customer networks are not advertised in the IGP at the provider edge router?

A) IGP performance will increase because the IGP will carry fewer routes.

B) Security is improved because the provider network will hide the customer network.

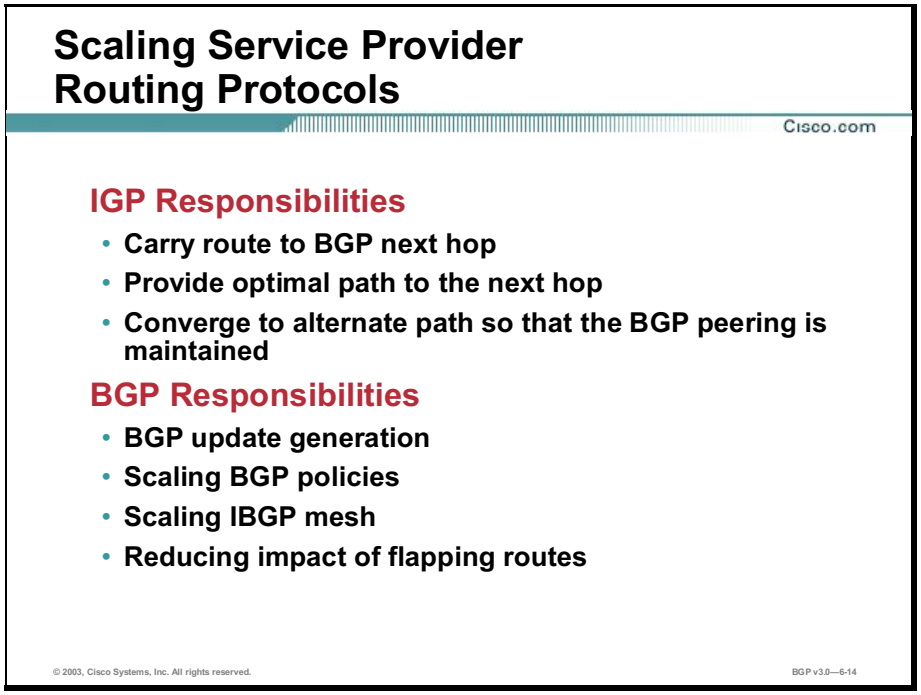
C) Route summarization is easier because the IGP will carry fewer routes.

D) A flapping link between the customer and the provider will not affect the stability of the provider IGP.

 Copyright © 2003, Cisco Systems, Inc.

Scaling Service Provider Routing Protocols

This topic describes proper scaling of IGPs and BGP in service provider networks.

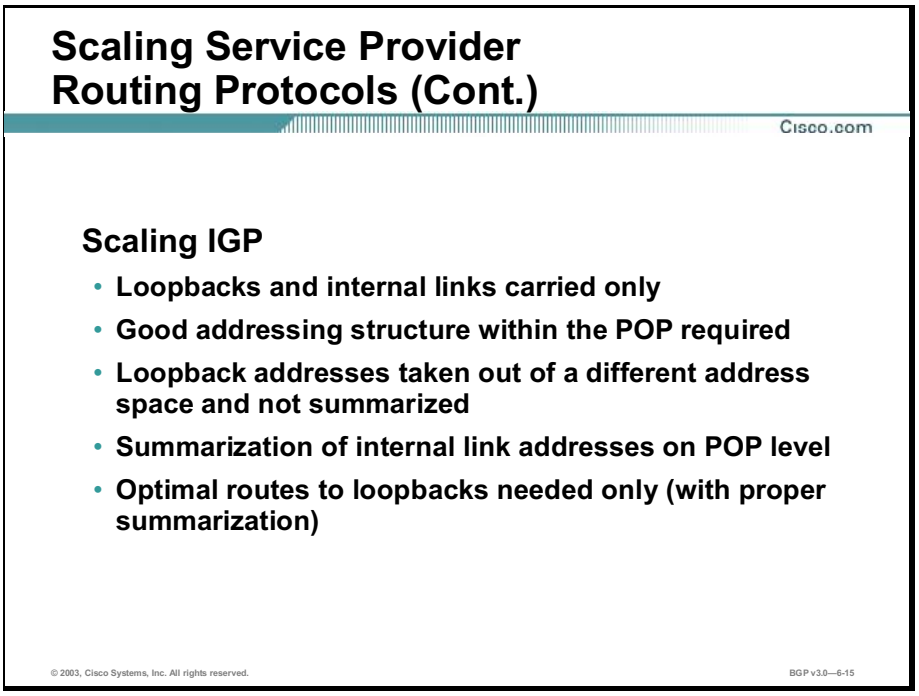

Scaling Service Provider Routing Protocols
Cisco.com
IGP Responsibilities
• Carry route to BGP next hop
• Provide optimal path to the next hop
• Converge to alternate path so that the BGP peering is maintained
BGP Responsibilities
• BGP update generation
• Scaling BGP policies
• Scaling IBGP mesh
• Reducing impact of flapping routes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-14

The IGP is responsible for:

- Carrying routes to the BGP next hops to facilitate recursive routing
- Providing an optimal path to the next hop, thereby optimizing packet flow toward all BGP destinations
- Converging to an alternate path in the case of lost links or routers in a redundant network; convergence should be quick so that BGP sessions are not lost

The BGP is responsible for:

- Generating BGP updates about reachable and unreachable networks
- Implementing and scaling the BGP routing policy, which can be quite cumbersome in large service provider networks with many EBGP-speaking routers
- Implementing and scaling IBGP sessions between all BGP-speaking routers in the AS
- Reducing the impact of individual flapping routes through route summarization


Scaling Service Provider Routing Protocols (Cont.)
Cisco.com
Scaling IGP
• Loopbacks and internal links carried only
• Good addressing structure within the POP required
• Loopback addresses taken out of a different address space and not summarized
• Summarization of internal link addresses on POP level
• Optimal routes to loopbacks needed only (with proper summarization)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-15

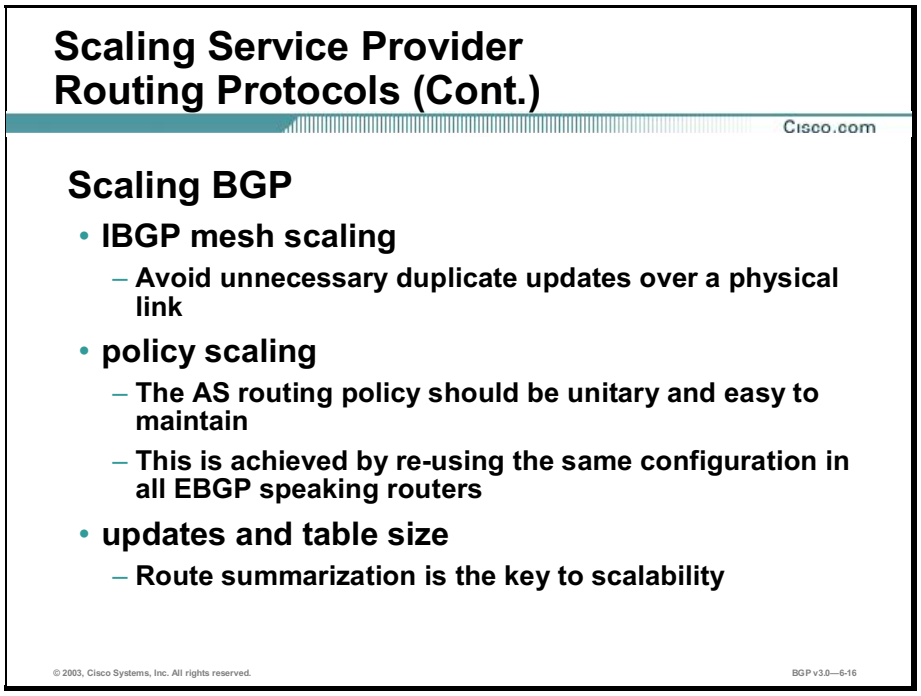
In scaling an IGP, it is important to limit the number of routes carried by the IGP. Optimally, the IGP carries only loopback interfaces and internal links.

The number of routes carried by the IGP can be even further reduced with route summarization. However, care must be taken because loopback addresses should never be summarized. Route summarization always introduces the risk of suboptimal routing and should be carefully planned, because it is important that recursive routing lookup always use optimal routing to the next hop. Also, in an MPLS environment, a label switched path (LSP) must be unbroken between edge routers, and summarizing loopback interfaces will break the LSP.

Internal links can always be summarized because they are not used as BGP next-hop addresses. To facilitate proper route summarization, internal links and loopback interfaces on a router should be assigned addresses from two different address spaces. Also, a router's internal links should be assigned addresses depending upon which POP they belong to.

If implemented correctly, all internal router links in one POP could be summarized at the POP level and injected into the core as a single route. But, all router loopback addresses within the POP are still propagated into the core as individual host routes, giving optimal routing to all loopback interfaces.

 Copyright © 2003, Cisco Systems, Inc.


Scaling Service Provider Routing Protocols (Cont.)
Cisco.com
Scaling BGP
• IBGP mesh scaling
– Avoid unnecessary duplicate updates over a physical link
• policy scaling
– The AS routing policy should be unitary and easy to maintain
– This is achieved by re-using the same configuration in all EBGP speaking routers
• updates and table size
– Route summarization is the key to scalability
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-16

The task of scaling BGP is actually three different and independent scaling tasks:

- BGP policy scaling. The AS routing policy should be unitary and easy to maintain. Different edge routers of the same AS should not use different policies and thereby advertise different routes to neighboring autonomous systems. Regardless of which router is currently active, the same routing policy should be in place. Administratively, replication of the same routing policies requires the same configuration lines in several edge routers.

- All BGP-speaking routers must be updated with consistent IBGP information. In the traditional BGP approach, ensuring consistent routing information was achieved by establishing a full mesh of IBGP sessions between all routers within the AS. An IBGP full mesh is certainly not scalable, and several tools are now available to achieve the same results without the full mesh.

- The number of routes in the routing table and the number of updates sent and received represent the third scaling task. Route summarization is the key to this scalability.

Practice

Q1) What are the two responsibilities of the IGP? (Choose two.)

A) carrying internal networks for optimal routing

B) BGP next-hop resolution

C) route propagation to external autonomous systems

D) exchanging routing information with customer networks

Q2) What are the two responsibilities of BGP? (Choose two.)

A) advertising router loopback interfaces for IBGP sessions

B) propagation of customer routes

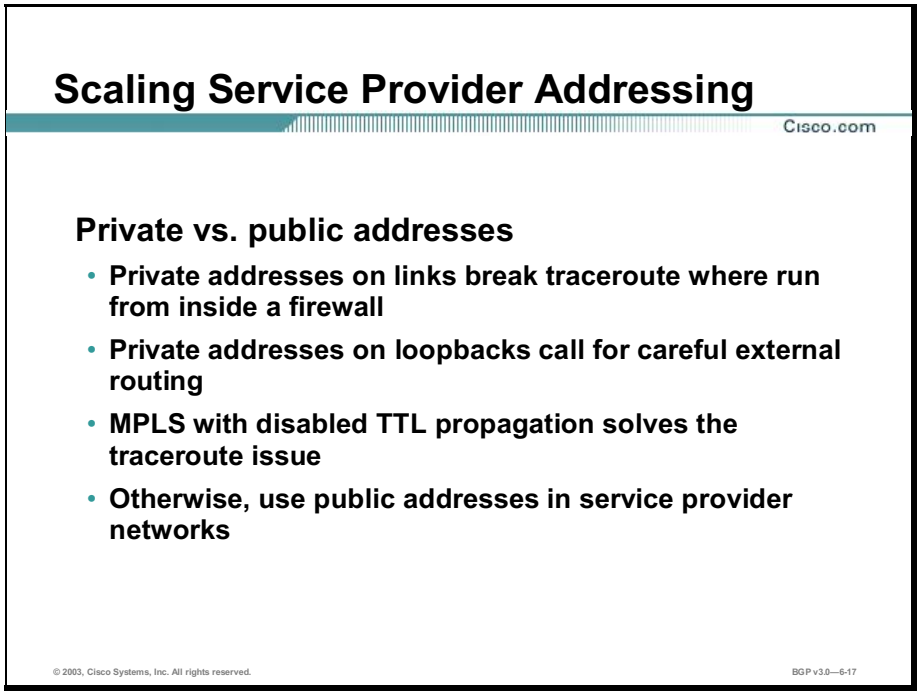
C) resolving BGP next-hop addresses

D) exchanging routing information with other providers

 Copyright © 2003, Cisco Systems, Inc.

Scaling Service Provider Addressing

This topic describes the scaling issues relevant to IP addressing in service provider networks.


Scaling Service Provider Addressing
Cisco.com
Private vs. public addresses
• Private addresses on links break traceroute where run from inside a firewall
• Private addresses on loopbacks call for careful external routing
• MPLS with disabled TTL propagation solves the traceroute issue
• Otherwise, use public addresses in service provider networks
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-17

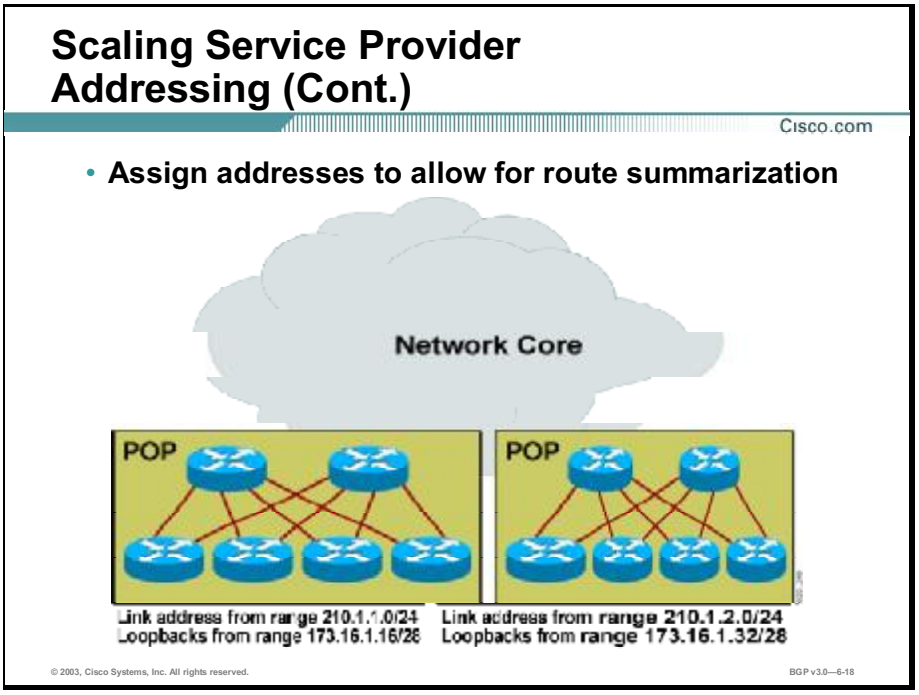
Using private addresses in a service provider network has some drawbacks. Private addresses on the provider internal links will cause trouble for the traceroute application. When the traceroute command is executed from a router inside a customer network that resides inside a firewall, the Internet Control Message Protocol (ICMP) replies generated by the provider router will have the source IP address assigned using the outgoing interface. If this is a private address, the customer firewall will most likely filter the packet due to address-spoofing detection rules. Even if the packet were allowed to enter the customer network, Domain Name System (DNS) reverse lookups would either fail or result in confusing printouts.

Using MPLS without Time to Live (TTL) propagation in the service provider network can easily overcome the traceroute problem with private addresses. If these functions are used, the provider network will appear as a single hop to the traceroute application. The intermediate routers will be invisible and thus can use private addresses.

Using private addresses on the service provider router loopback interfaces is possible. However, you must take care not to advertise any private addresses to any other AS.

A rule of safety is to prevent announcing any private addresses by using prefix-lists applied on outgoing updates to external neighbors. The same prefix-list mechanism can also be used on the provider edge routers to prevent accepting private addresses from any other AS if the other AS, by mistake, announces private addresses.

Example



In the figure, the left POP has been allocated two different address spaces. The address space 210.1.1.0/24 has been allocated to assign addresses to internal links within the POP. The address space 173.16.1.16/28 is used to assign addresses to loopback interfaces on routers within the POP.

Likewise, the right POP has assigned 210.1.2.0/24 to internal links and 173.16.1.32/28 to be used with loopback interfaces.

The two POPs connect to the core, and, as they do, both summarize the range for their internal links while they avoid summarizing the addresses assigned to the loopback interfaces of the POP router(s).

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What is the impact of using private addresses on the internal links of the service provider backbone?

A) If the provider uses private addressing, the customer must use public addressing.

B) A traceroute run from inside a customer network protected with a firewall is broken.

C) Private addressing prohibits optimal path selection in BGP.

D) If private addressing is used in the service provider backbone, route summarization must be disabled.

Q2) Why should loopback interface addresses not be summarized?

A) Summarizing the loopback interfaces will cause IBGP sessions to become inactive.

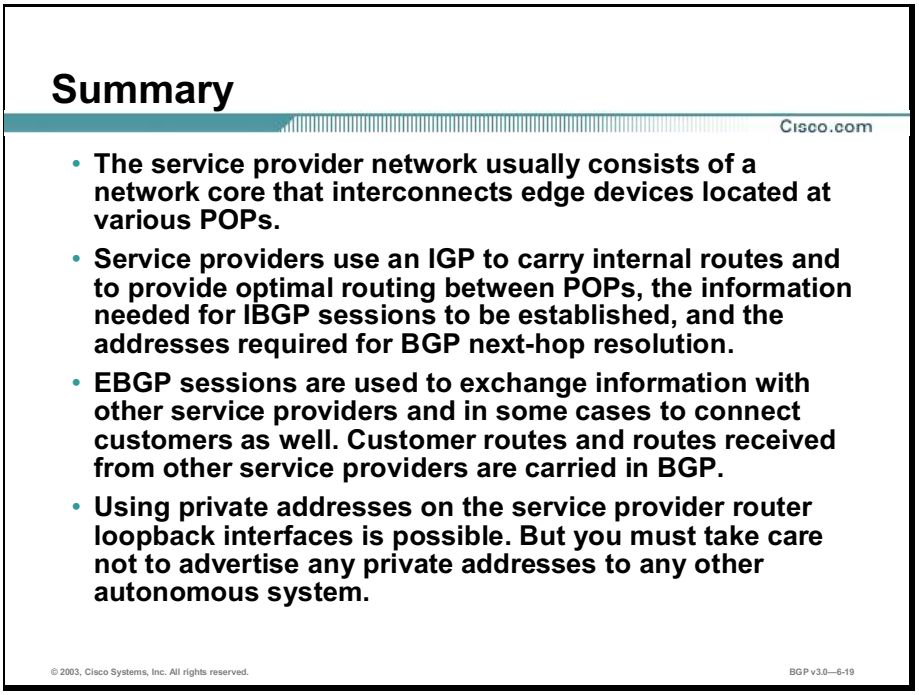
B) If the loopback interfaces are summarized, they will be unreachable from other routers in the POP.

C) Summarized loopback interfaces might interfere with optimal routing between loopback interfaces used for IBGP peering sessions.

D) Loopback interfaces can be summarized without any adverse effects.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• The service provider network usually consists of a network core that interconnects edge devices located at various POPs.
• Service providers use an IGP to carry internal routes and to provide optimal routing between POPs, the information needed for IBGP sessions to be established, and the addresses required for BGP next-hop resolution.
• EBGP sessions are used to exchange information with other service providers and in some cases to connect customers as well. Customer routes and routes received from other service providers are carried in BGP.
• Using private addresses on the service provider router loopback interfaces is possible. But you must take care not to advertise any private addresses to any other autonomous system.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-19

Next Steps

After completing this lesson, go to:

- Introduction to Route Reflectors lesson

References

For additional information, refer to these resources:

- For more information on scaling IGP and BGP in service provider networks, refer to "Designing Large-Scale IP Internetworks" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/idg4/nd2003.htm

 Copyright © 2003, Cisco Systems, Inc.

Quiz: Scaling IGP and BGP in Service Provider Networks

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe the basic structure of service provider networks
- Describe the propagation of internal and customer routes in service provider networks
- Describe proper scaling of IGPs and BGP in service provider networks
- Describe IP addressing scaling issues in service provider networks

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What three characteristics are common to typical service provider networks? (Choose three.)

A) The provider network uses two IGPs, one for customer routes and one for internal provider routes.

B) Service providers exchange routes with other providers using BGP.

C) Service providers run IBGP within their network in addition to their IGP requirements.

D) Service providers typically use either static routes or EBGP with their customers.

Q2) What is the typical role of an IGP within a service provider network?

A) The IGP carries customer routes for redistribution into BGP at the provider edge.

B) The IGP advertises a default route to customers of the service provider.

C) The IGP resolves next-hop IP addresses.

D) The IGP carries BGP routes across the provider network.

Q3) Why should you avoid the use of private IP addressing in service provider networks?

A) Private addressing can prevent customer network troubleshooting utilities such as traceroute from functioning correctly.

B) Private IP addressing is not allowed on the Internet and will not function in a service provider network.

C) Private IP addressing prevents the service provider from properly summarizing customer routes if they are also using private address space.

D) Private IP addressing prevents service provider applications such as MPLS from operating properly in an Internet-supporting environment.

Q4) What are three important implementation guidelines to properly scale BGP configurations in a service provider environment? (Choose three.)

A) IBGP full-mesh scaling tools to reduce duplicate traffic within the AS

B) summarization of customer routes to reduce the number of prefixes carried

C) improving BGP convergence time by using the IGP for route propagation within the provider AS

D) proper scaling of the AS-wide routing policy to ease administration and maintenance requirements

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 75 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Introduction to Route Reflectors

Overview

Two routers that have opened a Border Gateway Protocol (BGP) connection between them for the purpose of exchanging routing information are known as neighbors. Typically, routers within an autonomous system (AS) are configured in a full mesh. Route reflectors are a BGP scalability mechanism that enables routing information to be redistributed to all routers within an AS while eliminating the need for a fully meshed topology within the AS. This feature reduces the number of TCP sessions that must be maintained, lowering network overhead and CPU and memory resource requirements.

This lesson introduces BGP route reflectors by explaining why route reflectors improve BGP scalability. Modified split-horizon rules, when you are using route reflectors, are also discussed. The lesson concludes by describing the different redundancy mechanisms used with route reflectors, including route reflector clusters.

Importance

Large BGP networks cannot properly scale without relying on performance-enhancing tools such as route reflectors and confederations. Route reflectors enable BGP routing information to be distributed in a fashion that does not require a physical fully meshed network. Network overhead is reduced by decreasing the number of TCP connections required to distribute routing information and by lessening router CPU and memory requirements.

Objectives

Upon completing this lesson, you will be able to:

- Explain the need for BGP route reflectors
- Explain how route reflectors modify traditional IBGP split-horizon rules
- Explain the benefits of deploying redundant route reflectors

- Describe the concept of route reflector clusters
- Describe additional route reflector mechanisms designed to prevent routing loops

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- BGP Transit Autonomous Systems module

Outline

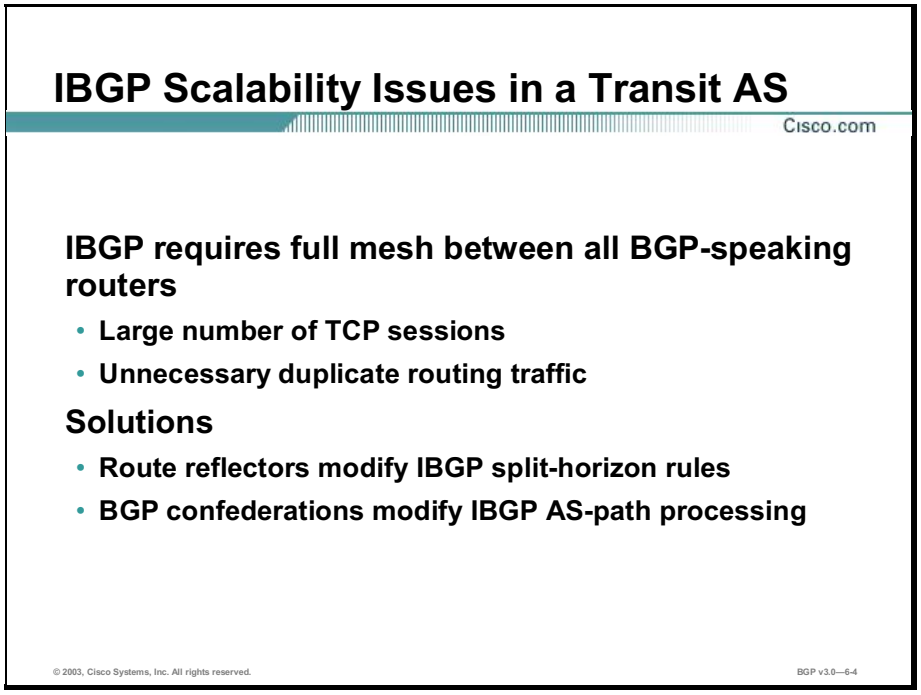
This lesson includes these topics:

- Overview
- IBGP Scalability Issues in a Transit AS
- Route Reflector Split-Horizon Rules
- Redundant Route Reflectors
- Route Reflector Clusters
- Additional Route Reflector Loop Prevention Mechanisms
- Summary
- Assessment (Quiz): Introduction to Route Reflectors

 Copyright © 2003, Cisco Systems, Inc.

IBGP Scalability Issues in a Transit AS

This topic explains the need for BGP route reflectors by describing the scalability issues of BGP transit backbones.


IBGP Scalability Issues in a Transit AS
Cisco.com
IBGP requires full mesh between all BGP-speaking routers
• Large number of TCP sessions
• Unnecessary duplicate routing traffic
Solutions
• Route reflectors modify IBGP split-horizon rules
• BGP confederations modify IBGP AS-path processing
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-4

Classic Internal Border Gateway Protocol (IBGP) split-horizon rules specify that updates received on an External Border Gateway Protocol (EBGP) session should be forwarded on all IBGP and EBGP sessions, but updates received on an IBGP session should be forwarded only to all EBGP sessions. This rule requires a BGP boundary router to be able to send routing updates to all other BGP-speaking routers in its own AS directly through a separate IBGP session to each of them.

The primary reason for the IBGP split-horizon rule is to avoid routing information loops within the AS. If the information received through an IBGP session is forwarded on other IBGP sessions, the information might come back to the originator, and be forwarded again in a never-ending loop. The originator would not detect the loop because no BGP attributes are changed on IBGP sessions.

The general design rule in classic IBGP is to have a full mesh of IBGP sessions. But a full mesh of IBGP sessions between “n” number of routers would require (n * (n-1)) / 2 IBGP sessions. For example, a router with an AS that contains 10 routers would require (10 * (10 – 1)) / 2 = 45 IBGP sessions. Imagine the number of sessions (and the associated router configuration) required for a single AS containing 500 routers.

Every IBGP session uses a single TCP session to another IBGP peer. An update that must be sent to all IBGP peers must be sent on each of the individual TCP sessions. If a router is attached to the rest of the network over just a single link, this single link has to carry all TCP/IP packets for all IBGP sessions. This requirement results in multiplication of the update over the single link.

Two different solutions are available to achieve greater scalability when you are faced with the full-mesh rules of IBGP autonomous systems:

- Route reflectors modify the classical IBGP split-horizon rules and allow a particular router to forward incoming IBGP updates to an outgoing IBGP session under certain conditions. This router becomes a concentration router, or a route reflector.

- BGP confederations (covered in a separate lesson) introduce the concept of a number of smaller autonomous systems within the original AS. The small autonomous systems exchange BGP updates between them using intraconfederation EBGP sessions.

Practice

Q1) What is a scalability limitation of IBGP-based transit autonomous systems?

A) A single AS is limited to 100 routers.

B) A full mesh of IBGP sessions is required due to BGP split-horizon rules.

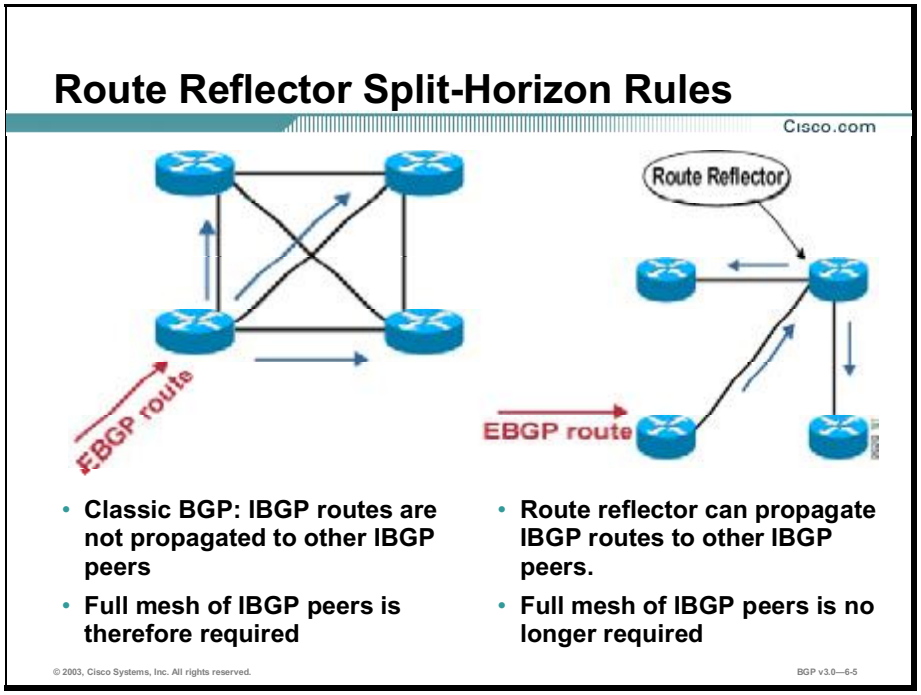
C) Only one instance of BGP is allowed in a single BGP domain.

D) Each BGP AS is limited to a single IGP.

 Copyright © 2003, Cisco Systems, Inc.

Route Reflector Split-Horizon Rules

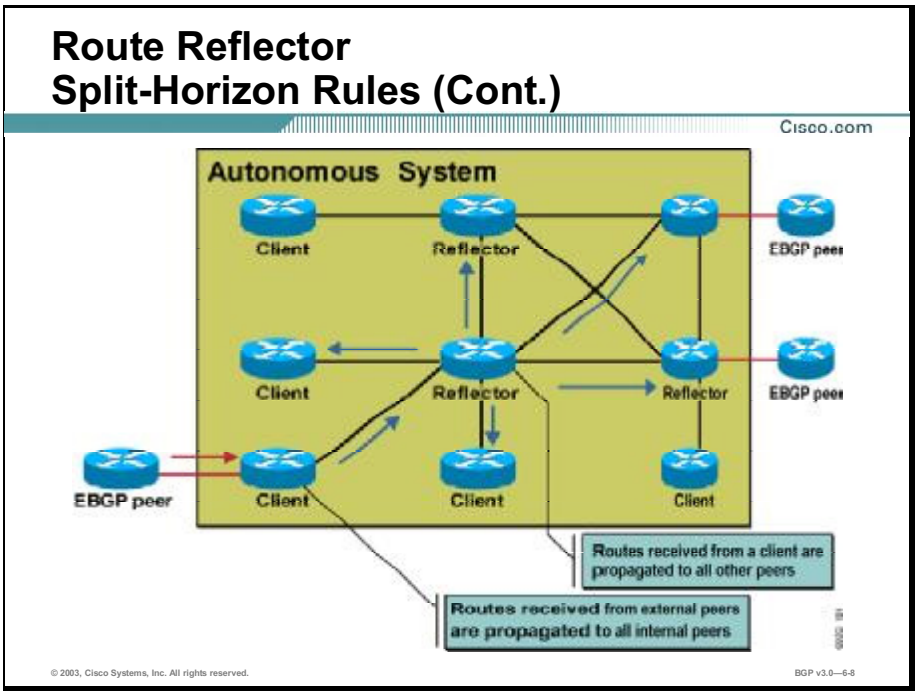
This topic explains how route reflectors modify traditional IBGP split-horizon rules.


Route Reflector Split-Horizon Rules
Cisco.com
Route Reflector
EBGP route
EBGP route
• Classic BGP: IBGP routes are not propagated to other IBGP peers
• Full mesh of IBGP peers is therefore required
• Route reflector can propagate IBGP routes to other IBGP peers.
• Full mesh of IBGP peers is no longer required
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-5

In classic IBGP, the BGP boundary router needs to forward the route received from an EBGP peer to each and every other router within its own AS using a dedicated IBGP session for each one. Also, the BGP boundary router forwards routes sourced by a router in the same way. To allow every router to update every other router, a full mesh of IBGP sessions is required.

The IBGP route reflector design relaxes the need for a full mesh. The router, configured as a route reflector, under certain conditions, will relay updates received through an IBGP session to another IBGP session. This capability requires modifications of the classic IBGP split-horizon rules.

The route reflector concept introduces processing overhead on the concentration router and, if configured incorrectly, can cause routing loops and instability.



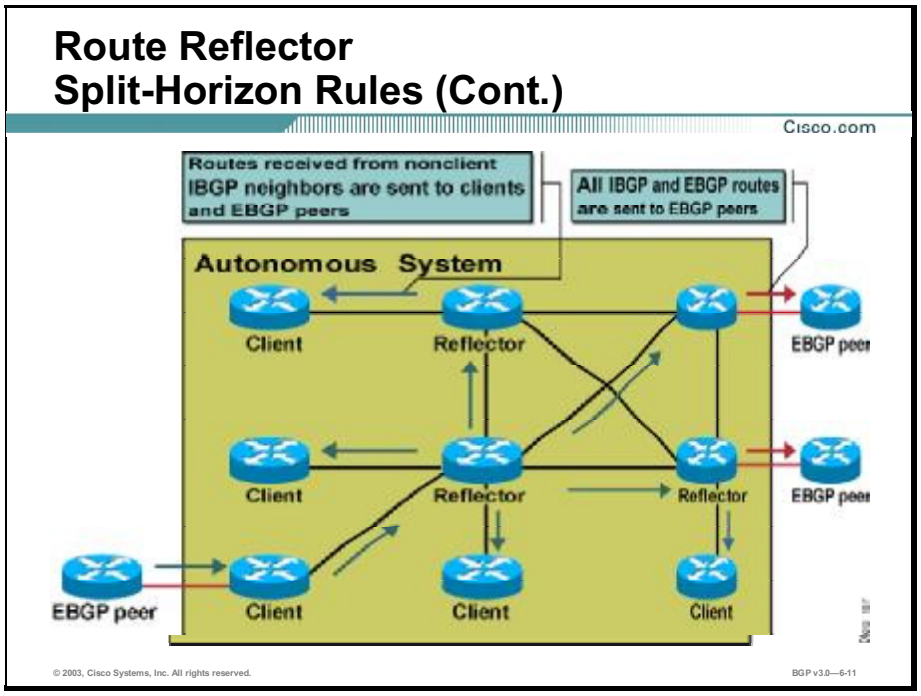
When you implement a route-reflector–based IBGP network, the BGP routers are divided into route reflectors (that implement modified split-horizon rules) and clients (that are behaving like traditional IBGP routers).

Route reflector clients are excluded from the full mesh. They can have any number of EBGP sessions but may have only one IBGP session, the session with their route reflector. Clients conform to the classic IBGP split-horizon rules and forward a received route from EBGP on its IBGP neighbor sessions. But the route reflector conforms to the route reflector split-horizon rules and recognizes that it has an IBGP session to a client. When the IBGP update is received from the client, the route reflector forwards the update to other IBGP neighbors, therefore alleviating the IBGP full-mesh requirement for its clients.

Similarly, when the route reflector receives an IBGP update from a neighbor that is not its client, it will forward the update to all of its clients.

Forwarding of an IBGP update in a route reflector does not change the next-hop attribute or any other common BGP attribute. This feature means that the client will use the most optimum route by means of recursive routing, regardless of the way that it has received the BGP route.

 Copyright © 2003, Cisco Systems, Inc.



The table presents detailed IBGP split-horizon rules as modified by the introduction of BGP route reflectors. For purposes of definition, a “route reflector” is a router that can perform the route reflection function. IBGP peers of the route reflector fall under two categories: “clients” and “nonclients.” The route reflector and its clients form a “cluster.” All IBGP peers of the route reflector that are not part of the cluster are nonclients. A “classic” IBGP router is a router that does not support route reflector functionality.

Type of router	Incoming update from	Is forwarded to
Classic	EBGP peer	All peers (IBGP and EBGP)
	IBGP peer	EBGP peers
Route reflector	EBGP peer	All peers (IBGP and EBGP)
	Nonclient IBGP peer	EBGP peers and clients
	Client IBGP peer	All peers but the sender
Client	EBGP peer	All peers (IBGP and EBGP)
	IGBP peer	EBGP peers

Practice

Q1) Which three types of routes does a route reflector propagate to its clients? (Choose three.)

A) EBGP

B) client IBGP

C) nonclient IBGP

D) IGP

Q2) Which routes are propagated by a route reflector client to its IBGP neighbors?

A) EBGP

B) client IBGP

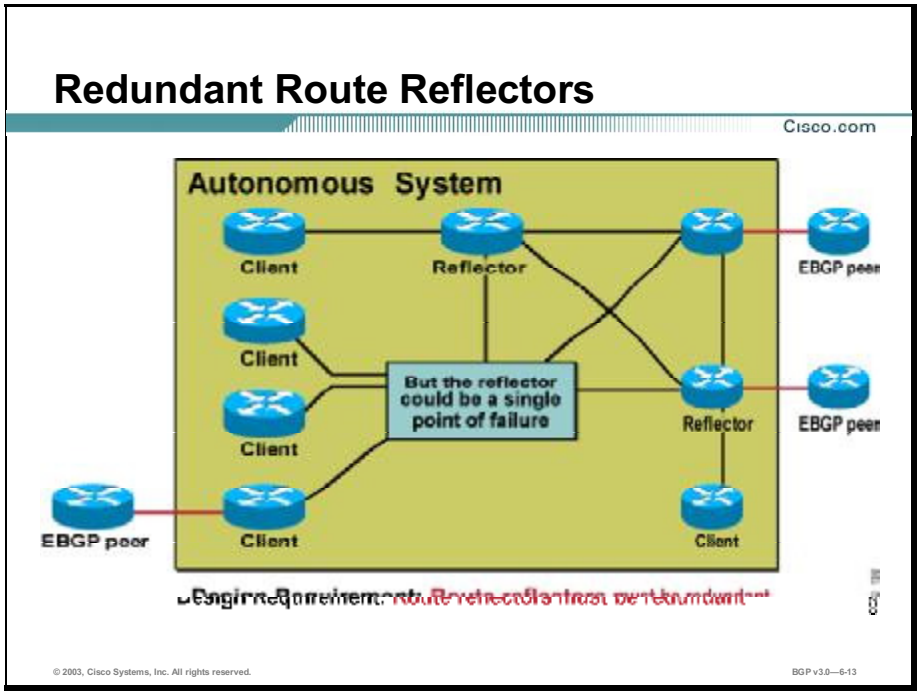
C) nonclient IBGP

D) IGP

 Copyright © 2003, Cisco Systems, Inc.

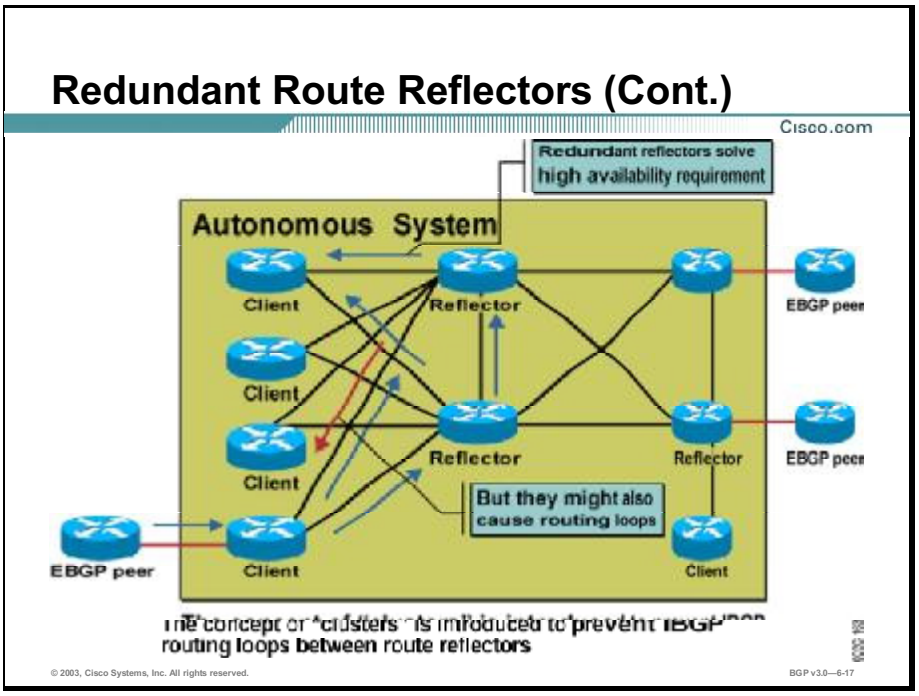
Redundant Route Reflectors

This topic explains the benefits of deploying redundant route reflectors.



Clients may have any number of EBGP peers but may have IBGP sessions only with their route reflector(s). If the reflector fails, its client can no longer send or receive BGP updates from the rest of the AS. The route reflector is, therefore, a single point of failure.

To avoid introducing a single point of failure into the network, the route reflector functionality must be as redundant as the physical network. If a client will still be physically attached to the network after its route reflector has failed, the client should have a redundant route reflector. Thus, in all highly available networks, route reflectors must be redundant.



A client may have IBGP sessions to more than one route reflector to avoid a single point of failure. Each client will receive the same route from all of its reflectors. Both route reflectors will receive the same IBGP update from their client, and they will both reflect the update to the rest of the clients. Additionally, both route reflectors will get updated from the full mesh and reflect those updates to their clients. As a result, each client will get two copies of all routes. Under certain circumstances (particularly when you use weights on IBGP sessions to influence BGP route selection), improper route reflection can result in an IBGP routing loop that is impossible to detect. Additional BGP attributes are thus necessary to prevent these routing loops.

Practice

Q1) What are a benefit and a drawback of redundant route reflector designs? (Choose two.)

A) Each client will receive only a single copy of BGP route update information.

B) IBGP routing loops can occur.

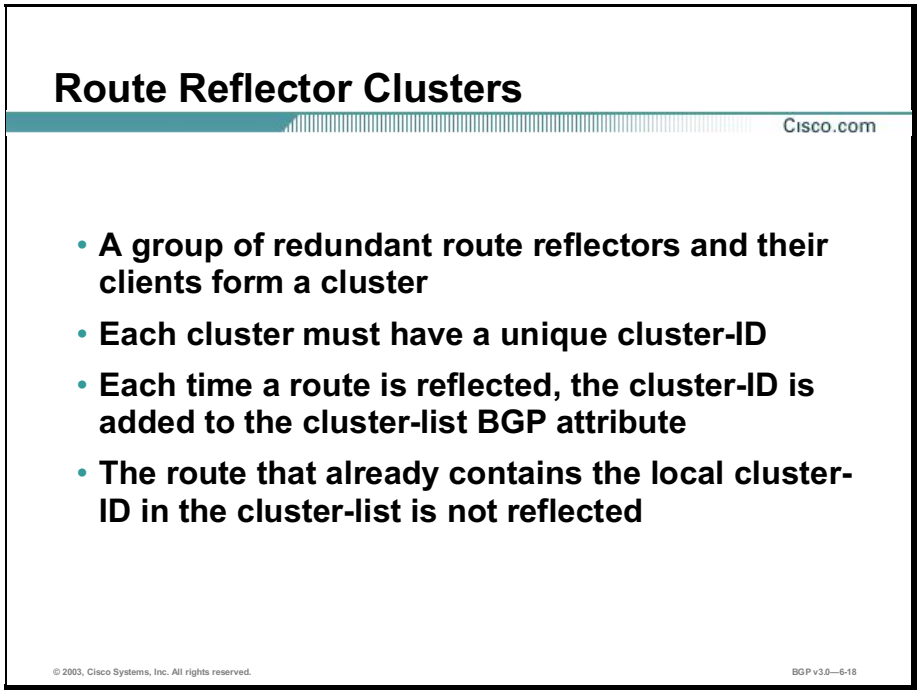
C) The transit AS does not need an IGP.

D) Redundancy eliminates a single point of failure in the route reflector.

 Copyright © 2003, Cisco Systems, Inc.

Route Reflector Clusters

This topic describes the concept of route reflector clusters and their loop prevention benefits when you are deploying route reflectors in redundant configurations.


Route Reflector Clusters
Cisco.com
• A group of redundant route reflectors and their clients form a cluster
• Each cluster must have a unique cluster-ID
• Each time a route is reflected, the cluster-ID is added to the cluster-list BGP attribute
• The route that already contains the local cluster-ID in the cluster-list is not reflected
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-18

A router acting as a route reflector client does not require any specific configuration. It simply has fewer IBGP sessions than it would have if it were part of the full mesh. But improperly configuring the client to also be a reflector could easily cause a loop. An IBGP route coming in from one of the real reflectors to the client could be forwarded by the client, erroneously acting as reflector, to the other reflector.

Route reflector clusters prevent IBGP routing loops in redundant route reflector designs.

The role of the network designer is to properly identify which route reflectors and their clients will form a cluster. The designer assigns to the cluster a cluster-ID number that is unique within the AS.

Note	The cluster-ID number must be configured in the route reflectors. The clients should not be configured with this information.

A route reflector router can reflect routes only within a single cluster. A route reflector can, however, participate in another cluster but only as a client. A client can function as a client only to a route reflector(s) belonging to the same cluster.

When a route is reflected, the reflector creates the cluster-list attribute and attaches it to the route if it does not already exist. It then sets its cluster-ID number in the cluster-list or adds its cluster-ID number to an already existing cluster-list attribute. If the route, for any reason, is ever reflected back to the same reflector, it will recognize its cluster-ID number in the cluster-

list and not forward it again. The first route reflector that reflects the route also sets an additional BGP attribute, called “originator-ID,” and adds it to the BGP router-ID of its client.

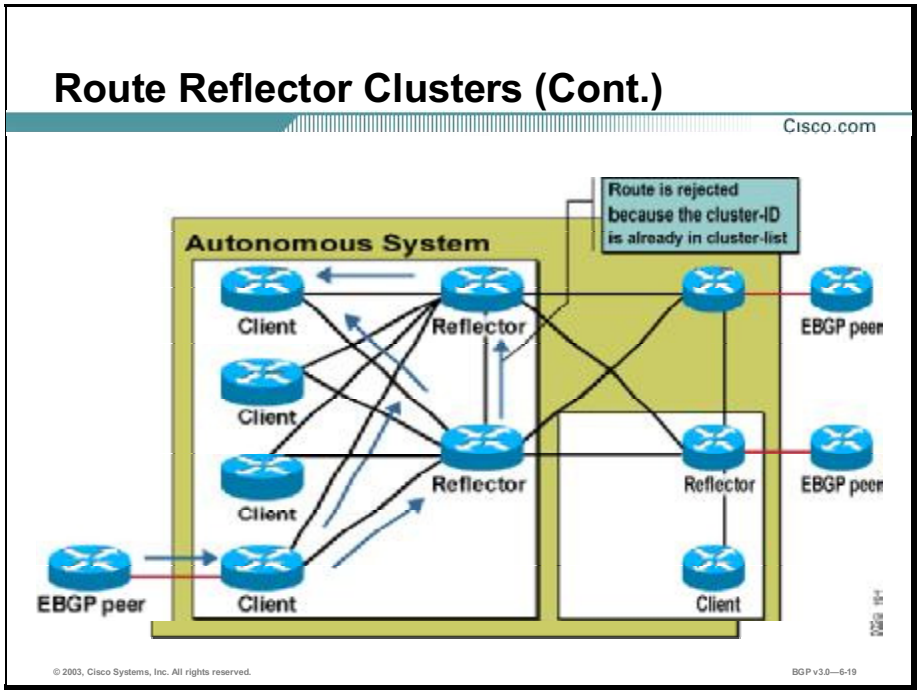
Note The cluster-list and originator-ID attributes are nontransitive optional BGP attributes, allowing routers that do not support route reflector functionality to coexist with route reflectors and their clients in the same AS.

Based on cluster-list and originator-ID attributes, routers can implement two loop prevention mechanisms:

- Any router that receives an IBGP update with the originator-ID attribute set to its own BGP router-ID will ignore that update.
- Any route reflector that receives an IBGP update with its cluster-ID already in the cluster-list will ignore that update.

 Copyright © 2003, Cisco Systems, Inc.

Example



In a cluster with redundant route reflectors, the client will forward the received EBGP update to both reflectors. The route reflectors forward the update into the IBGP full mesh. This behavior means that they will send the update to each other as well. But when a route reflector receives a BGP update from another route reflector, it will recognize their common cluster-ID number in the cluster-list attribute. Therefore, the newly received route update is ignored.

Practice

Q1) What are route reflector clusters used for?

A) to detect and correct route instability

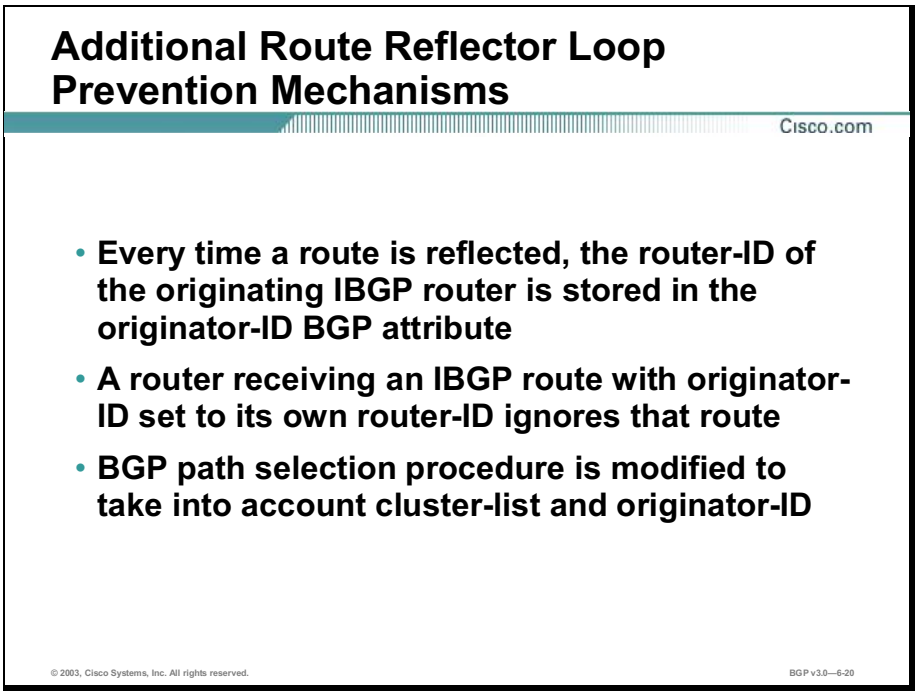
B) to differentiate between route reflector clients and nonclients

C) to eliminate EBGP split-horizon rules

D) to detect potential routing loops in redundant configurations

Additional Route Reflector Loop Prevention Mechanisms

This topic describes additional route reflector mechanisms designed to prevent routing loops.


Additional Route Reflector Loop Prevention Mechanisms
Cisco.com
• Every time a route is reflected, the router-ID of the originating IBGP router is stored in the originator-ID BGP attribute
• A router receiving an IBGP route with originator-ID set to its own router-ID ignores that route
• BGP path selection procedure is modified to take into account cluster-list and originator-ID
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-20

When a route is reflected, the route reflector will set the originator-ID BGP attribute (nontransitive optional BGP attribute) to the router-ID of the peer from which it received the route. Any router receiving a route with its own router-ID in the originator-ID attribute silently ignores that route.

BGP path selection rules have been modified to select the best route in scenarios where a router might receive reflected and nonreflected routes or several reflected routes:

- The traditional BGP path selection parameters (weight, local preference, origin, multi-exit discriminator [MED]) are compared first.
- If these parameters are equal, the routes received from EBGP neighbors are preferred over routes received from IBGP neighbors.
- When a router receives two IBGP routes, the nonreflected routes (routes with no originator-ID attribute) are preferred over reflected routes.
- The reflected routes with shorter cluster-lists are preferred over routes with longer cluster-lists.
- If the additional route-reflector–oriented selection criteria do not yield a decision, the rest of the traditional BGP path selection rules are followed.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What two loop-prevention mechanisms support route reflectors? (Choose two.)

A) cluster-list

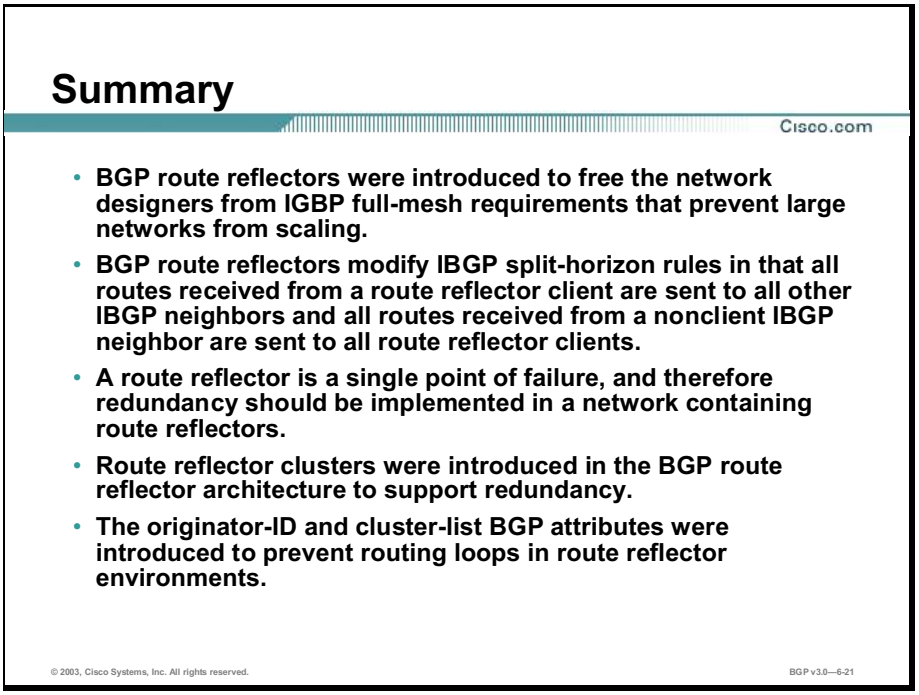
B) originator-ID

C) AS-path attribute

D) spanning tree

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• BGP route reflectors were introduced to free the network designers from IGBP full-mesh requirements that prevent large networks from scaling.
• BGP route reflectors modify IBGP split-horizon rules in that all routes received from a route reflector client are sent to all other IBGP neighbors and all routes received from a nonclient IBGP neighbor are sent to all route reflector clients.
• A route reflector is a single point of failure, and therefore redundancy should be implemented in a network containing route reflectors.
• Route reflector clusters were introduced in the BGP route reflector architecture to support redundancy.
• The originator-ID and cluster-list BGP attributes were introduced to prevent routing loops in route reflector environments.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-21

Next Steps

After completing this lesson, go to:

- Network Design with Route Reflectors lesson

References

For additional information, refer to these resources:

- For more information on route reflectors, refer to “BGP Case Studies Section 4” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#routereflectors

 Copyright © 2003, Cisco Systems, Inc.

Quiz: Introduction to Route Reflectors

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Explain the need for BGP route reflectors
- Explain how route reflectors modify traditional IBGP split-horizon rules
- Explain the benefits of deploying redundant route reflectors
- Describe the concept of route reflector clusters
- Describe additional route reflector mechanisms designed to prevent routing loops

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What is the main problem solved by implementing BGP route reflectors?

A) the large number of routes carried in the IGP when BGP is deployed

B) the ability for BGP to scale a single AS in a large network

C) the need for a homogeneous method of applying policies to routes carried through an AS

D) the lack of a service level agreement supporting features and the need to deploy these features with greater ease

Copyright © 2003, Cisco Systems, Inc.

Q1) How does a route reflector modify the IBGP split-horizon rule?

A) Route reflectors forward EBGP updates onto all peers (IBGP and EBGP).

B) Route reflectors treat all neighbors as EBGP peers, eliminating the IBGP mesh requirements.

C) Route reflectors forward IBGP updates from clients to other IBGP neighbors.

D) Route reflectors append the cluster-ID to the AS path, allowing peers to be treated as EBGP neighbors.

Q2) Why are redundant route reflectors mandatory in any high-availability network design?

A) because all neighbors peer with the route reflector, and a large number of neighbors can make the route reflector router unstable

B) because EBGP peers can inject BGP updates into the AS only through the route reflector

C) because route reflectors maintain more routing information, making them more prone to congestion and failure

D) because clients can form IBGP relationships only with the route reflector

Q1) What is the main reason for implementing redundant route reflectors with clusters?

A) to eliminate routing loops in redundant configurations

B) to limit the number of neighbor sessions with each route reflector

C) to provide another scalability mechanism targeted at removing the IBGP full-mesh requirement

D) to enhance security within the AS

 Copyright © 2003, Cisco Systems, Inc.

Q2) How does the originator-ID attribute assist in the elimination of routing loops caused by redundant route reflector designs?

A) If the originator-ID matches the router-ID of the reflector, local preference is set on the route to make it a backup.

B) The originator-ID attribute is set to the cluster-ID to ensure that a route traverses the AS only one time.

C) A router receiving a route with the originator-ID matching its router-ID will ignore that route.

D) The originator-ID allows the router to know if the route originated locally or from an external source so that administrative distance rules for the route can be verified.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Network Design with Route Reflectors

Overview

Route reflectors are a Border Gateway Protocol (BGP) scalability mechanism that enables routing information to be redistributed to all routers within an autonomous system (AS) while eliminating the need for a fully meshed topology within the AS. Properly implementing these features requires careful network design within the AS.

This lesson introduces the network design rules that network designers should follow when implementing a network with BGP route reflectors. It also lists the potential issues that can arise if the network design rules are not adhered to. The lesson concludes by describing the concept of hierarchical route reflectors.

Importance

Large BGP networks cannot properly scale without relying on performance-enhancing tools such as route reflectors and confederations. Route reflectors enable BGP routing information to be distributed in a fashion that does not require a physical full-mesh network. Implementing such a network requires careful design, or routing issues such as lost routes or loops can occur.

Objectives

Upon completing this lesson, you will be able to:

- List the network design rules for implementing BGP route reflectors
- List the potential issues that can arise if you do not follow the route reflector network design rules
- Explain the purpose and function of hierarchical route reflectors

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- BGP Transit Autonomous Systems module

Outline

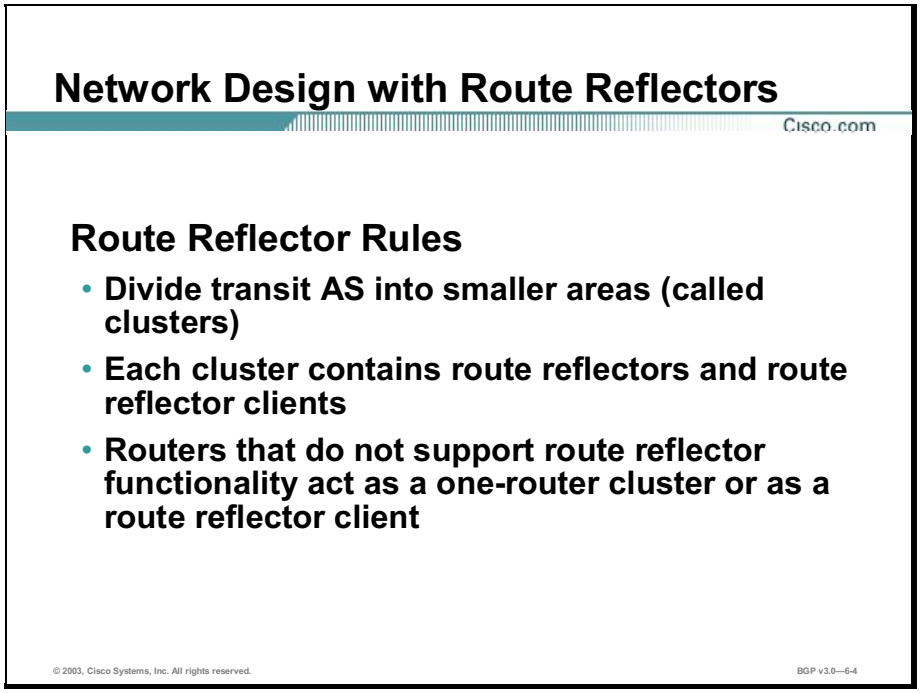
This lesson includes these topics:

- Overview
- Network Design with Route Reflectors
- Potential Network Issues
- Hierarchical Route Reflectors
- Summary
- Assessment (Quiz): Network Design with Route Reflectors

 Copyright © 2003, Cisco Systems, Inc.

Network Design with Route Reflectors

This topic lists the network design rules for implementing BGP route reflectors. Included in this topic are the design rules for route reflectors and Internal Border Gateway Protocol (IBGP) sessions.


Network Design with Route Reflectors
Cisco.com
Route Reflector Rules
• Divide transit AS into smaller areas (called clusters)
• Each cluster contains route reflectors and route reflector clients
• Routers that do not support route reflector functionality act as a one-router cluster or as a route reflector client
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-4

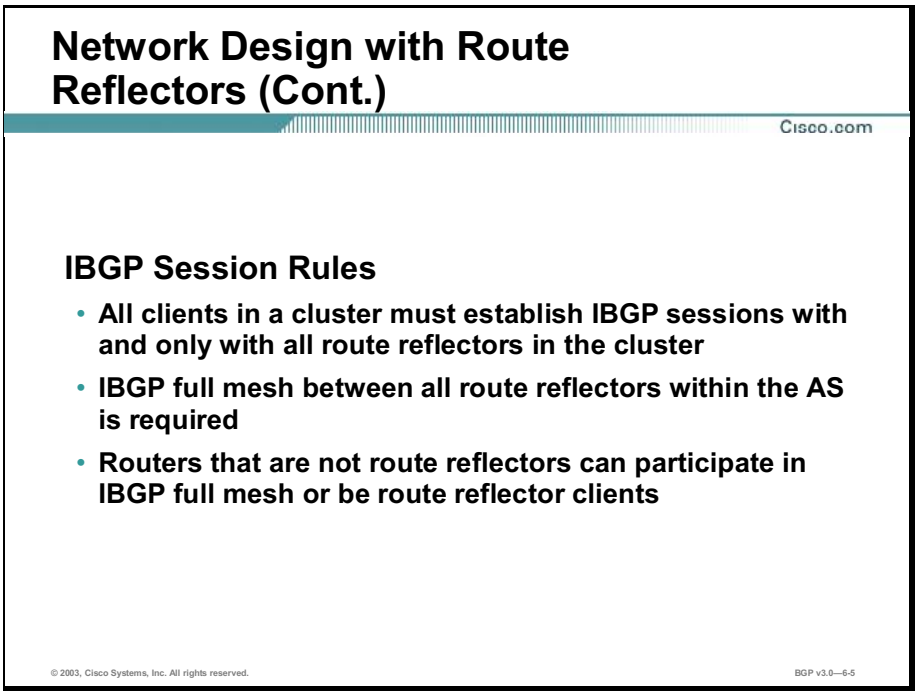
The physical topology of the network could serve as a guide to route reflector design.

Implementing route reflectors within the transit AS will create smaller areas (or groups) of routers. These smaller groupings of routers are called clusters. A cluster consists of route reflector routers, either redundant or nonredundant, and the client routers connected to them.

In designing the implementation of route reflectors within a transit AS, identify a group of peripheral routers that are physically connected to the same backbone router or routers. Consider the peripheral routers as clients and the backbone routers as route reflectors. Then, consider this group of routers together to form a cluster.

Note	Additional design examples and rules are available in the BGP Transit Autonomous System module.

Only the routers configured as route reflectors require a Cisco IOS® software version with route reflector functionality. A router lacking this functionality in its installed Cisco IOS software can function as a client or be a part of the full mesh.


Network Design with Route Reflectors (Cont.)
Cisco.com
IBGP Session Rules
• All clients in a cluster must establish IBGP sessions with and only with all route reflectors in the cluster
• IBGP full mesh between all route reflectors within the AS is required
• Routers that are not route reflectors can participate in IBGP full mesh or be route reflector clients
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-5

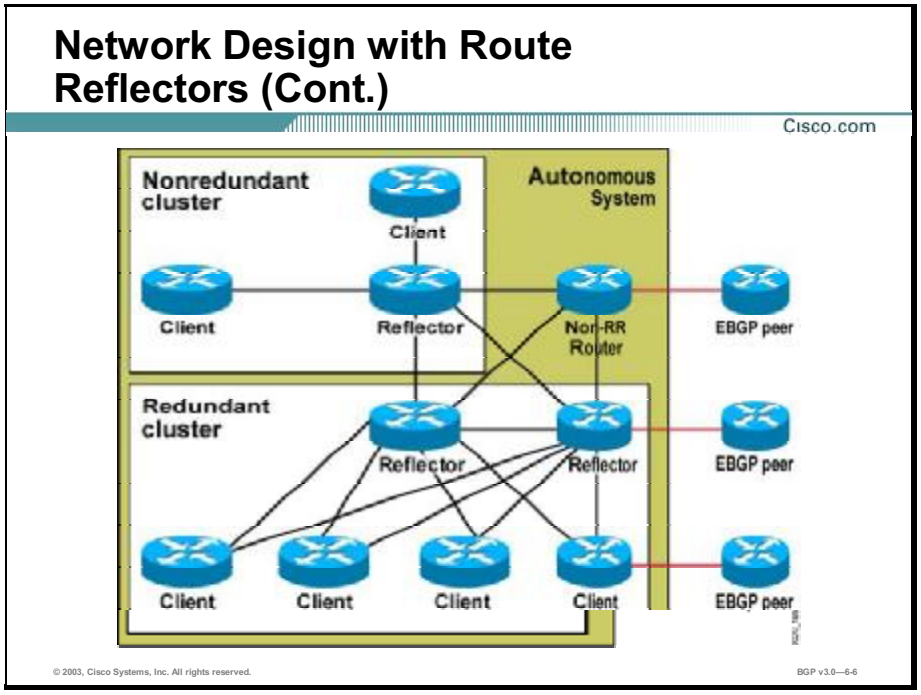
The principal goal for designing networks with BGP route reflectors is to reduce the size of the full mesh of IBGP sessions by excluding some routers from the mesh. The routers excluded from the full mesh, the clients, have to send their IBGP information to, and receive it from, at least one router that belongs to the full mesh, the route reflector. Thus, the full mesh is still there, but it is smaller, and all route reflectors have to be part of it.

All clients in a cluster should have IBGP sessions with all their route reflectors and their route reflectors only. If a client does not have sessions with all the reflectors in the cluster, the redundancy is violated. If a client has IBGP sessions to routers other than the route reflectors, unnecessary routing traffic will be generated.

Both clients and other routers, those that are not route reflectors, obey the classic IBGP split-horizon rules. Thus, non-route-reflector routers are either a client to a reflector or are participating directly in the full mesh.

 Copyright © 2003, Cisco Systems, Inc.

Example



In this example, the routers serving as route reflectors and the non-route-reflector router have IBGP sessions in a full mesh.

In the area called the “redundant cluster,” the four client routers and the two route reflector routers make up the cluster. Each of the four client routers has an IBGP session with the two route reflectors and only with those two route reflectors.

In the nonredundant area, each of the two client routers has a single physical connection to a route reflector router. These three routers form a nonredundant cluster. The router designated as the route reflector in the cluster is already a single point of failure in this physical design because a failure of this router will prevent the clients in the cluster from reaching the rest of the network. Therefore, there is no new single point of failure introduced when the router is configured as the only route reflector in this cluster. Each of the two clients has a single IBGP session to the route reflector.

The other router shown is not configured as a route reflector nor is it a client to any other route reflector. It serves as an example of where a non-route-reflector router participates in the full mesh.

Practice

Q1) With what should a route reflector client have IBGP sessions?

A) with other clients only

B) with its closest neighbor EBGP peer only

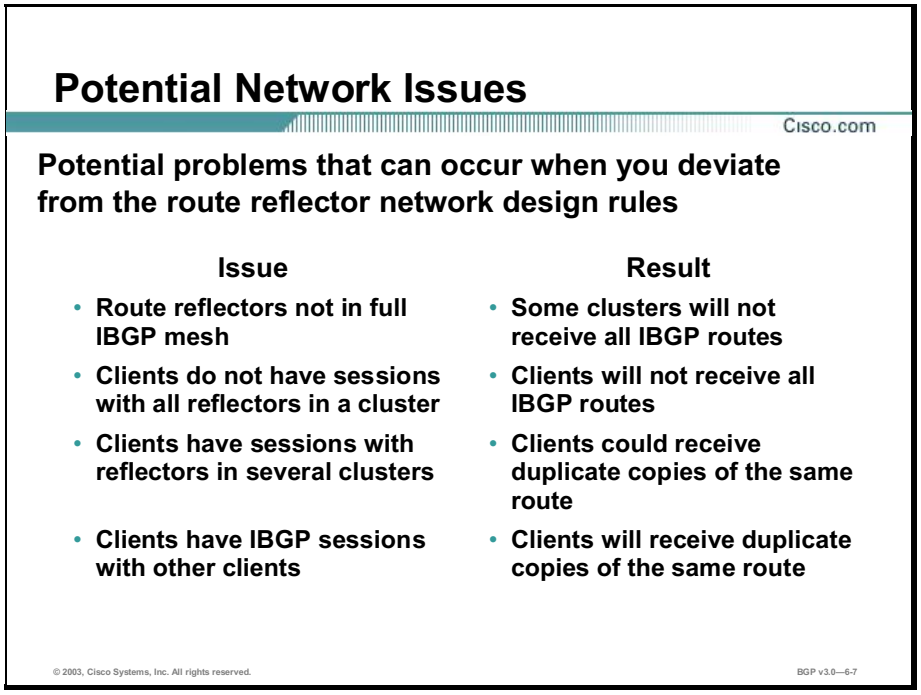
C) with its route reflectors and directly connected EBGP peers

D) with its route reflectors and other clients in the cluster

 Copyright © 2003, Cisco Systems, Inc.

Potential Network Issues

This topic lists the potential issues that can arise if the route reflector network design rules as explained in the previous topic are not followed.


Potential Network Issues
Cisco.com
Potential problems that can occur when you deviate from the route reflector network design rules
Issue
Route reflectors not in full IBGP mesh
Clients do not have sessions with all reflectors in a cluster
Clients have sessions with reflectors in several clusters
Clients have IBGP sessions with other clients
Result
Some clusters will not receive all IBGP routes
Clients will not receive all IBGP routes
Clients could receive duplicate copies of the same route
Clients will receive duplicate copies of the same route
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-7

Two nontransitive optional BGP attributes, originator-ID and cluster-list, are both used to prevent fatal loops of information. The use of these two attributes makes a network fairly insensitive to poor configuration, however. Thus, for optimal performance, you must have an optimal configuration. Here are some of the problems that could occur if you deviate from route reflector network design rules:

- If route reflectors are not connected with IBGP sessions in a full mesh, some clusters will not have all the routes.
- If a client has IBGP sessions with some route reflectors in a cluster, but not all of them, the client might miss some BGP routes.
- If a client has IBGP sessions to route reflectors that belong to different clusters, the BGP update from the client will be forwarded by the client into the full mesh with different cluster-IDs in the cluster-list attribute. When the BGP update enters the mesh, it will reach the other route reflector, which will, unnecessarily, accept the route as valid and forward it into its cluster. This situation, in turn, causes unnecessary duplication of updates to the clients.
- If a client has IBGP sessions to other clients in the same cluster, those clients will receive unnecessary duplications of updates.

Practice

Q1) What happens if a route reflector client establishes sessions with route reflectors in two clusters?

A) This is an invalid configuration.

B) It introduces the originator-ID attribute and the potential for routing loops.

C) BGP updates from the client will not be forwarded into the full mesh with different cluster-IDs.

D) Some routes will be reflected by more than one cluster, resulting in increased size of the BGP table for route reflectors and for the offending client.

Q2) What happens if two route reflectors have each other configured as a client?

A) This is an invalid configuration.

B) Routing loops will occur because the cluster-list will not be updated properly.

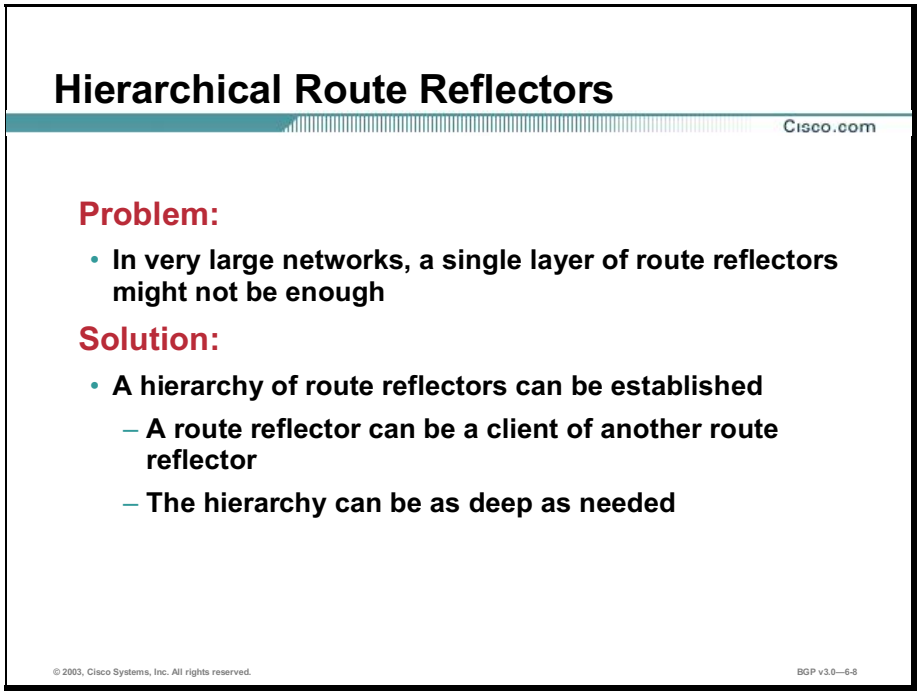
C) The only result will be increased BGP traffic.

D) BGP updates from the client will not be forwarded into the full mesh with different cluster-IDs.

 Copyright © 2003, Cisco Systems, Inc.

Hierarchical Route Reflectors

This topic explains the purpose and function of hierarchical route reflectors.


Hierarchical Route Reflectors
Cisco.com
Problem:
• In very large networks, a single layer of route reflectors might not be enough
Solution:
• A hierarchy of route reflectors can be established
– A route reflector can be a client of another route reflector
– The hierarchy can be as deep as needed
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-8

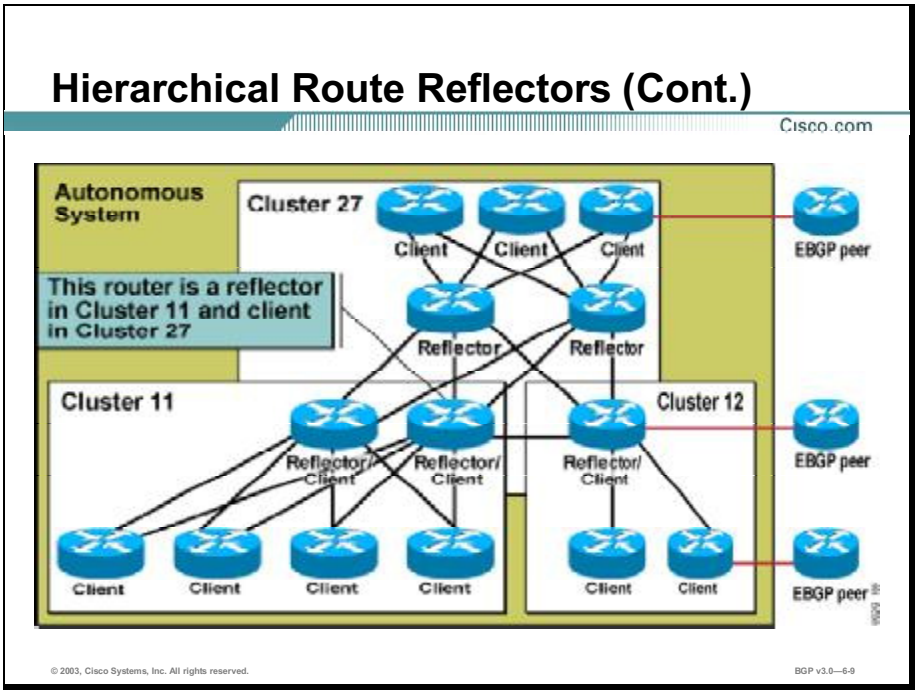
Network designers can build route reflector clusters in hierarchies. With hierarchies, a router serving as a route reflector in one cluster can act as a client in another cluster.

Clients are not configured to be route reflector clients; they simply have fewer IBGP sessions. On the contrary, a network designer must configure a route reflector. In configuring an IBGP session on a route reflector, the designer must configure the session to reach a client for the route reflector IBGP split-horizon rules to start working. All other IBGP sessions configured on the route reflector are a part of the full mesh. Also, the designer must configure the cluster-ID on the route reflector.

A router that is configured to be a route reflector will still have ordinary IBGP sessions that are part of the full mesh. If these sessions are reduced in number and only a few remain, and the remaining ones reach a second level of route reflectors, a hierarchy of route reflectors is created.

When a designer builds a first level of clusters, the remaining full mesh is smaller than when all routers belonged to it. But if it is large enough, the designer can build an additional level of route reflectors.

Example



In this example, the first level of route reflector clusters was built by creating cluster 11 and cluster 12. This first step reduced the original full mesh of 14 routers to a full mesh of 8 routers.

A second level of route reflector clusters was built by creating cluster 27. This second step further reduced the full mesh of eight routers to a full mesh consisting of only two routers. Only the two route reflectors in cluster 27 should be connected in a full mesh.

When a client in the lowest level receives an EBGP update, it will forward it on all configured IBGP sessions to a route reflector. The route reflector recognizes BGP updates received from configured clients and will forward these updates to all other clients using normal IBGP sessions. The update, sent on a normal IBGP session, will be a second-level client update to the second-level route reflector. The second-level route reflector will recognize that the update was received from a client and will forward it to all other clients and into the full mesh.

Practice

Q1) What is a hierarchical route reflector design?

A) a design where there is only one route reflector per cluster

B) a design where route reflectors are located in distinct, nonoverlapping clusters

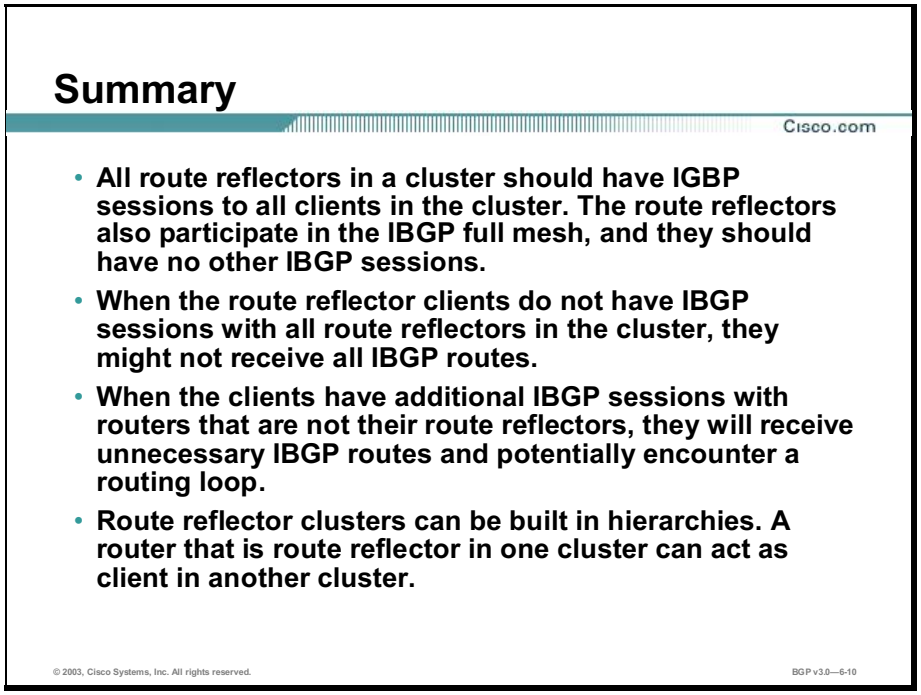
C) a design where some route reflectors are also clients of other route reflectors

D) a design with multiple route reflectors per cluster

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• All route reflectors in a cluster should have IGBP sessions to all clients in the cluster. The route reflectors also participate in the IBGP full mesh, and they should have no other IBGP sessions.
• When the route reflector clients do not have IBGP sessions with all route reflectors in the cluster, they might not receive all IBGP routes.
• When the clients have additional IBGP sessions with routers that are not their route reflectors, they will receive unnecessary IBGP routes and potentially encounter a routing loop.
• Route reflector clusters can be built in hierarchies. A router that is route reflector in one cluster can act as client in another cluster.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-10

Next Steps

After completing this lesson, go to:

- Configuring and Monitoring Route Reflectors lesson

References

For additional information, refer to these resources:

- For more information on route reflectors, refer to “BGP Case Studies Section 4” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#routereflectors

Quiz: Network Design with Route Reflectors

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- List the network design rules for implementing BGP route reflectors
- List the potential issues that can arise if you do not follow the route reflector network design rules
- Explain the purpose and function of hierarchical route reflectors

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What can occur if a client has IBGP neighbor relationships with other routers not configured as route reflectors?

A) This is an invalid configuration.

B) The client will notify the route reflector and be promoted to a route reflector as well.

C) Routing black holes can occur, causing lost traffic inside the AS.

D) Unnecessary routing traffic will be generated.

 Copyright © 2003, Cisco Systems, Inc.

Q2) What potential problem can occur if a client does not have an IBGP session with all route reflectors in a cluster?

A) This is an invalid configuration.

B) The client might not receive all BGP routes.

C) EBGP routes received by the client will not be distributed properly throughout the AS.

D) Duplicate routing traffic will be sent to the client.

Q3) What problem are hierarchical route reflectors designed to solve?

A) lack of a consistent application of security and routing policies throughout the AS

B) scalability of autonomous systems in very large routing domains

C) routing loops caused by redundant cluster configurations

D) administrative overhead when you are implementing router reflector network designs

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 100 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Configuring and Monitoring Route Reflectors

Overview

This lesson introduces the steps required to successfully migrate an existing autonomous system (AS) to Border Gateway Protocol (BGP) route reflectors. It also lists the Cisco IOS® commands required to configure and monitor route reflectors.

Importance

Large BGP networks cannot properly scale without relying on performance-enhancing tools such as route reflectors and confederations. Route reflectors enable BGP routing information to be distributed in a fashion that does not require a physical full-mesh network. Implementing such a network requires knowledge of the steps to properly migrate and configure route reflectors, and the commands used to verify the operation of a configured network.

Objectives

Upon completing this lesson, you will be able to:

- List the steps to migrate an existing IBGP backbone to a backbone with route reflectors
- Identify the Cisco IOS commands required to configure route reflectors on a BGP backbone
- Identify the Cisco IOS commands required to monitor a BGP backbone containing route reflectors

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- BGP Transit Autonomous Systems module

Outline

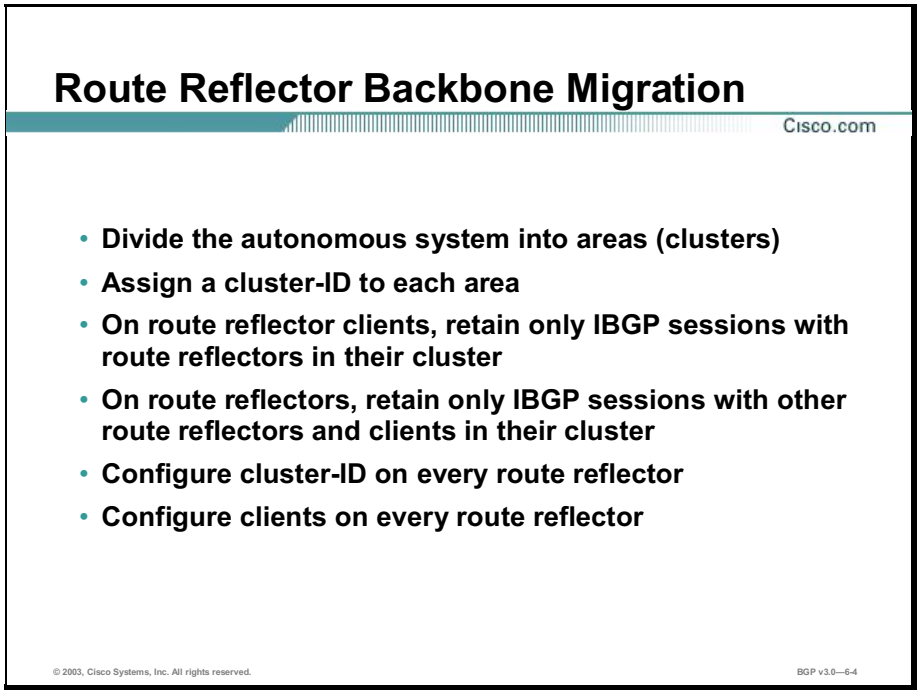
This lesson includes these topics:

- Overview
- Route Reflector Backbone Migration
- Configuring Route Reflectors
- Monitoring Route Reflectors
- Summary
- Assessment (Lab): BGP Route Reflectors

 Copyright © 2003, Cisco Systems, Inc.

Route Reflector Backbone Migration

This topic lists the steps required to successfully migrate an existing IBGP backbone to a backbone with route reflectors.


Route Reflector Backbone Migration
Cisco.com
• Divide the autonomous system into areas (clusters)
• Assign a cluster-ID to each area
• On route reflector clients, retain only IBGP sessions with route reflectors in their cluster
• On route reflectors, retain only IBGP sessions with other route reflectors and clients in their cluster
• Configure cluster-ID on every route reflector
• Configure clients on every route reflector
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-4

The physical topology of the AS serves as a guide to design clusters. You should introduce no additional single points of failure when you are deploying route reflectors. If the physical topology is redundant, a good practice is to have redundant route reflectors. If the physical topology is not redundant, introducing a nonredundant cluster does not add a single point of failure because the network was already nonredundant.

The following planning and preparation steps are required before migrating from a full mesh of IBGP sessions to a route reflector design.

Step 1 Identify a group of peripheral routers that are physically connected to the same set of backbone routers. Consider the peripheral routers as clients and the backbone routers as route reflectors. Let the routers form a cluster. Make sure that no router belongs to two different clusters, because this would represent an illegal configuration.

Step 2 Create a numbering plan that indicates how numbers are assigned to the clusters in the network. The plan must make sure to uniquely identify each of the clusters within the AS. Clusters are not seen from outside the AS, so the plan does not need to be coordinated with any other AS. To ease troubleshooting, it is recommended that numbers lower than 256 be used, because cluster-IDs are displayed in IP address format.

Note The default value of a cluster-ID is the BGP router-ID of the route reflector. If you decide to implement nonredundant clusters, you do not have to plan the cluster-ID numbers, because the BGP router-IDs should be unique.

Practice

Q1) Which two BGP parameters do you have to configure on a route reflector? (Choose two.)

A) cluster-ID

B) originator-ID

C) cluster-list

D) route reflector clients

Q2) Which BGP parameter do you have to configure on a route reflector client?

A) cluster-ID

B) originator-ID

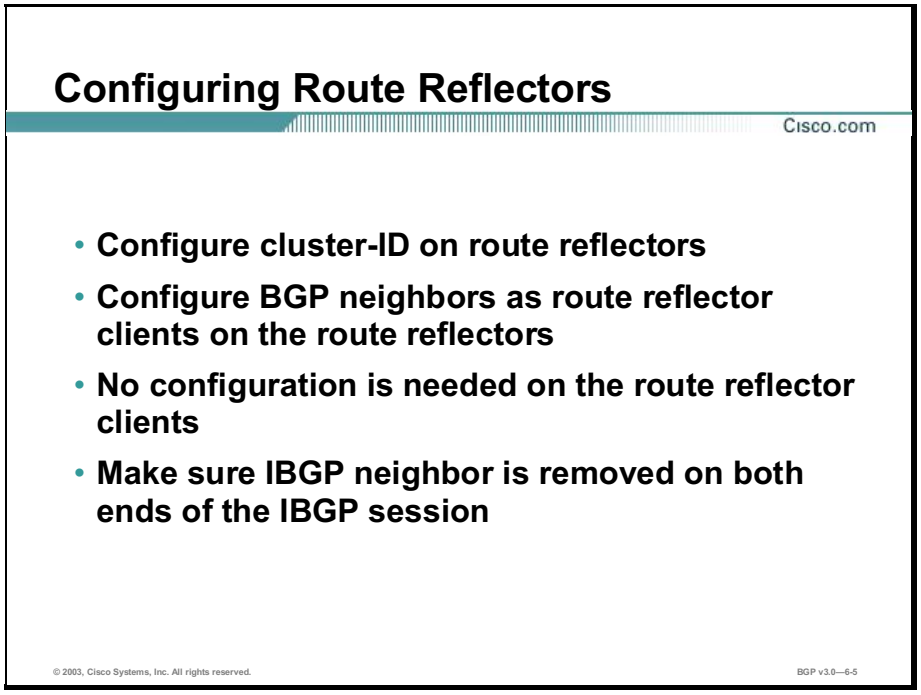
C) route reflector identification

D) No configuration is required on the client routers.

 Copyright © 2003, Cisco Systems, Inc.

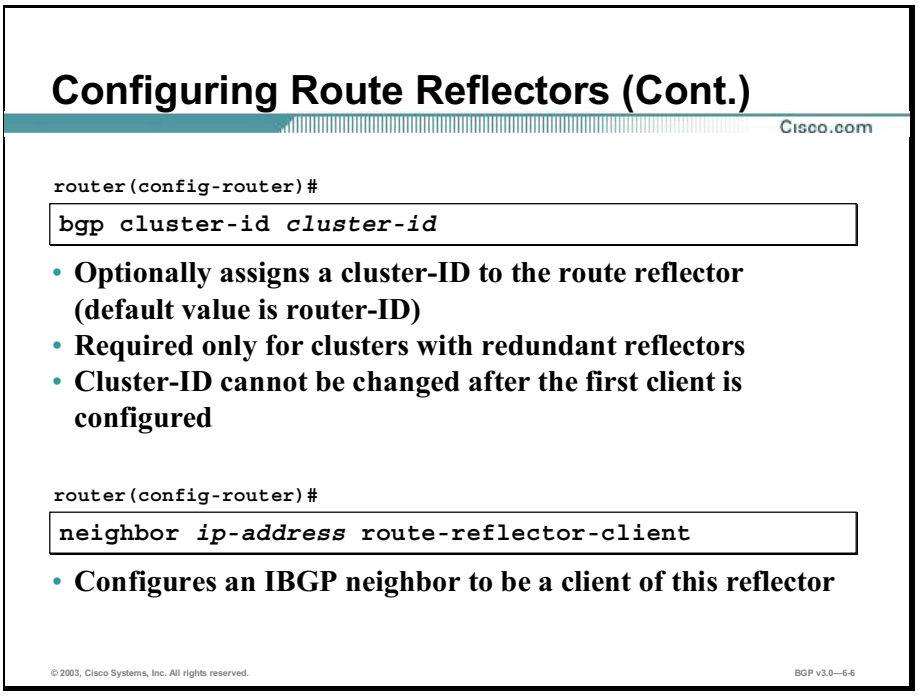
Configuring Route Reflectors

This topic lists the steps and the Cisco IOS commands required to configure BGP route reflectors.


Configuring Route Reflectors
Cisco.com
• Configure cluster-ID on route reflectors
• Configure BGP neighbors as route reflector clients on the route reflectors
• No configuration is needed on the route reflector clients
• Make sure IBGP neighbor is removed on both ends of the IBGP session
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-5

As part of the planning and preparation necessary to migrate from a full mesh of IBGP sessions to a route reflector design, you need to make the following configuration changes:

- Configure the proper cluster-ID value on the route reflectors.
- Configure the route reflector with information about which IBGP neighbor sessions are reaching their clients.
- In the clients, remove all IBGP sessions to neighbors that are not a route reflector in the client cluster.
- Make sure that the IBGP neighbor is removed on both ends of the IBGP session.


Configuring Route Reflectors (Cont.)
Cisco.com
router(config-router)#
bgp cluster-id cluster-id
• Optionally assigns a cluster-ID to the route reflector (default value is router-ID)
• Required only for clusters with redundant reflectors
• Cluster-ID cannot be changed after the first client is configured
router(config-router)#
neighbor ip-address route-reflector-client
• Configures an IBGP neighbor to be a client of this reflector
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-6

bgp cluster-id

Command used to configure the cluster-ID if the BGP cluster has redundant route reflectors.

bgp cluster-id *cluster-id*

To remove the cluster-ID, use the **no** form of this command.

no bgp cluster-id *cluster-id*

Syntax Description

cluster-id	Cluster-ID of this router acting as a route reflector; the cluster-ID is a maximum of 4 bytes.

 Copyright © 2003, Cisco Systems, Inc.

neighbor route-reflector-client

Command used to configure the router as a BGP route reflector and configure the specified neighbor as its client. When all the clients are disabled, the local router is no longer a route reflector.

neighbor *ip-address* **route-reflector-client**

To indicate that the neighbor is not a client, use the **no** form of this command.

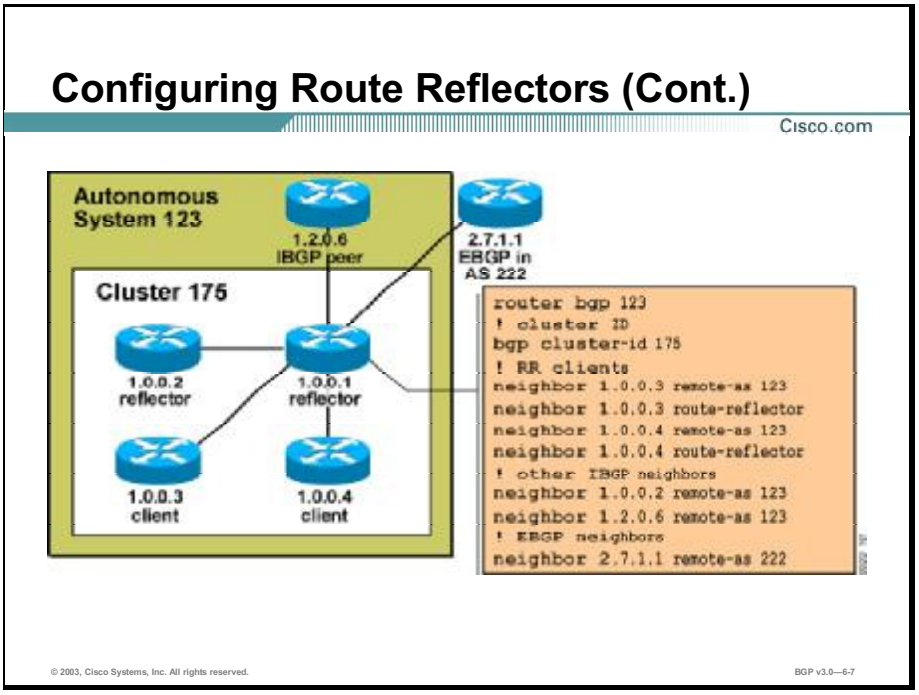
no neighbor *ip-address* **route-reflector-client**

Syntax Description

ip address	Neighbor IP address

By default, there is no route reflector in the AS.

Example



In this example, AS 123 has been divided into clusters with route reflectors. The routers with router-ID 1.0.0.1 and 1.0.0.2 are route reflectors in a cluster that has been assigned cluster-ID 175. The routers with router-ID 1.0.0.3 and 1.0.0.4 (full command would be **neighbor** *ip-address* **route-reflector-client**) are clients to these two route reflectors.

The figure shows a portion of the configuration in router 1.0.0.1. The cluster-ID is assigned to the router under the **router bgp** process definition of the router configuration. After being assigned, the route reflector client configuration is added under the **router bgp** process for the two neighbors identifying the two sessions reaching clients.

Practice

Q1) What are three migration steps required to convert from a fully meshed IBGP AS to an AS based on route reflectors? (Choose three.)

A) Remove unnecessary IBGP sessions.

B) Configure the clients on the route reflectors.

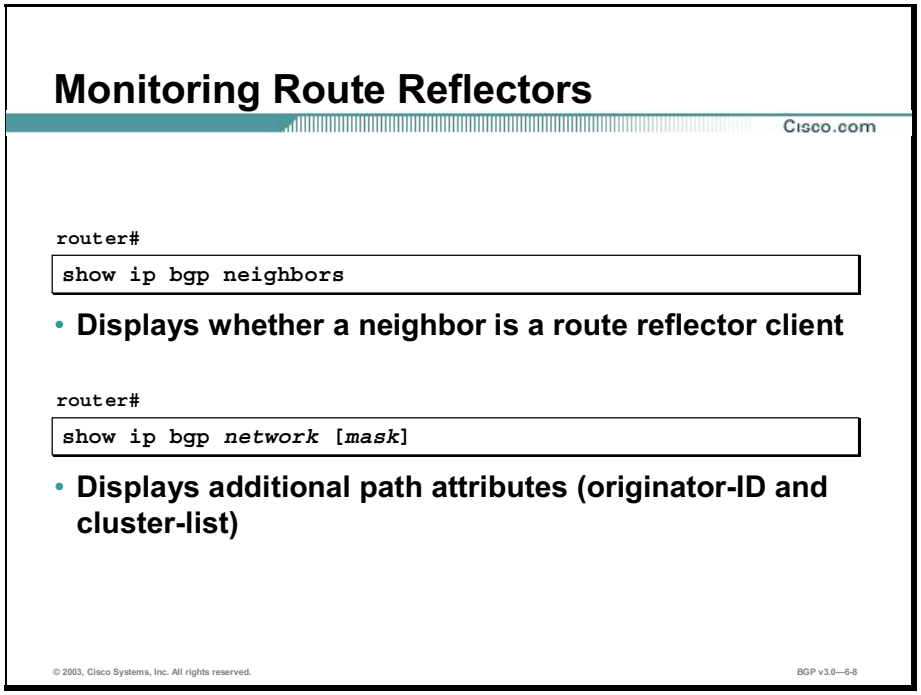
C) Configure IBGP sessions between route reflector clients.

D) Configure the cluster-ID on the route reflectors.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring Route Reflectors

This topic lists the Cisco IOS commands required to monitor route reflector configurations.


Monitoring Route Reflectors
Cisco.com
router#
show ip bgp neighbors
• Displays whether a neighbor is a route reflector client
router#
show ip bgp network [mask]
• Displays additional path attributes (originator-ID and cluster-list)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-8

show ip bgp neighbors

To display information about the TCP and BGP connections to neighbors, use the **show ip bgp neighbors** EXEC command.

show ip bgp neighbors [*address*] [**received-routes** | **routes** | **advertised-routes** | {**paths** *regular-expression*} | **dampened-routes**]

In this case, the **show ip bgp neighbors** command is used on the router, not to see routes or paths received but to see the status of the neighbor session, so no other qualifiers than the optional IP address are given.

show ip bgp

To display entries in the BGP routing table, use the **show ip bgp** EXEC command.

show ip bgp [*network*] [*network-mask*] [**longer-prefixes**]

When details are displayed for a specific route entry in the BGP table, the cluster-list and originator-ID attributes are also shown.

Monitoring Route Reflectors (Cont.)

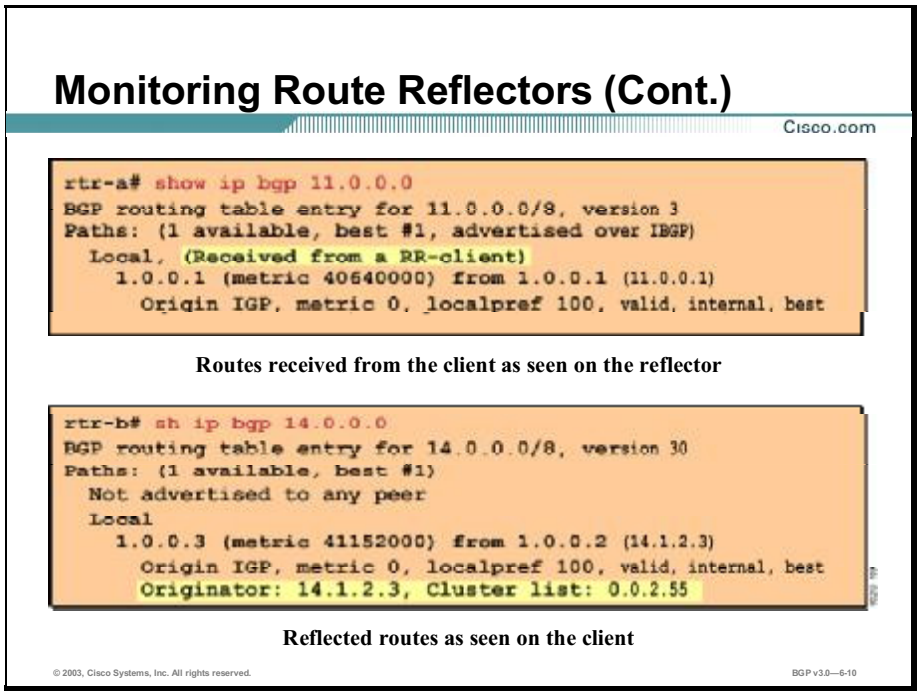
Cisco.com

```
router# show ip bgp neighbors 1.0.0.1
BGP neighbor is 1.0.0.1, remote AS 213, internal link
 Index 1, Offset 0, Mask 0x2
  Route-Reflector Client
  BGP version 4, remote router ID 11.0.0.1
  BGP state = Established, table version = 5, up for 01:33:24
  Last read 00:00:24, hold time is 180, keepalive interval is
60 seconds
  Minimum time between advertisement runs is 5 seconds
  Received 257 messages, 0 notifications, 0 in queue
  Sent 264 messages, 0 notifications, 0 in queue
  Connections established 5; dropped 4
  Last reset 01:33:33, due to : User reset request
  No. of prefix received 1
```

© 2003, Cisco Systems, Inc. All rights reserved.

The **show ip bgp neighbors** command, issued on the route reflector router, indicates that the neighbor is a route reflector client.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Route Reflectors (Cont.)
Cisco.com
rtr-a# show ip bgp 11.0.0.0
BGP routing table entry for 11.0.0.0/8, version 3
Paths: (1 available, best #1, advertised over IBGP)
Local, (Received from a RR-client)
1.0.0.1 (metric 40640000) from 1.0.0.1 (11.0.0.1)
Origin IGP, metric 0, localpref 100, valid, internal, best
Routes received from the client as seen on the reflector
rtr-b# sh ip bgp 14.0.0.0
BGP routing table entry for 14.0.0.0/8, version 30
Paths: (1 available, best #1)
Not advertised to any peer
Local
1.0.0.3 (metric 41152000) from 1.0.0.2 (14.1.2.3)
Origin IGP, metric 0, localpref 100, valid, internal, best
Originator: 14.1.2.3, Cluster list: 0.0.2.55
Reflected routes as seen on the client
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-10

The first example of a **show ip bgp** command is issued on a route reflector router. It shows that this particular entry in the BGP table was received from a route reflector client.

The second example shows an entry in the BGP table that at some point was reflected from a route reflector. The reflecting router has added the originator-ID and cluster-list attributes to the route.

Practice

Q1) Which command should you use to display the cluster-list attribute, identifying whether an IBGP route was reflected?

A) **show ip bgp** *prefix*

B) **show ip bgp** *cluster*

C) **show ip bgp** *interface*

D) **show ip bgp** *attribute-id*

Q2) Which command should you use to identify route reflector clients without inspecting the router configuration?

A) **show ip bgp** *prefix*

B) **show ip bgp neighbors**

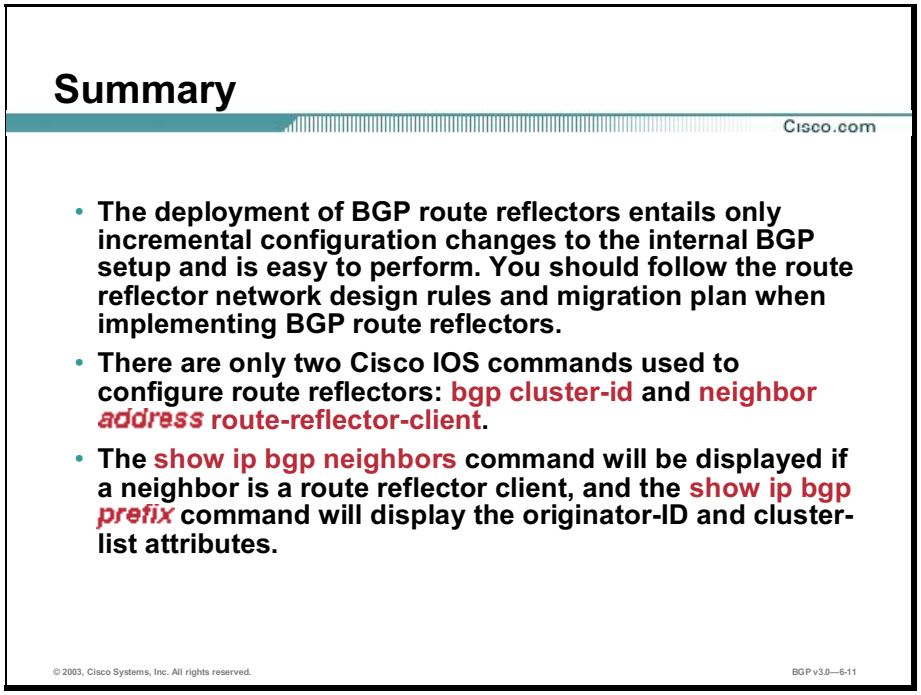
C) **show ip bgp clients**

D) **show ip bgp summary**

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• The deployment of BGP route reflectors entails only incremental configuration changes to the internal BGP setup and is easy to perform. You should follow the route reflector network design rules and migration plan when implementing BGP route reflectors.
• There are only two Cisco IOS commands used to configure route reflectors: bgp cluster-id and neighbor address route-reflector-client.
• The show ip bgp neighbors command will be displayed if a neighbor is a route reflector client, and the show ip bgp prefix command will display the originator-ID and cluster-list attributes.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-11

Next Steps

After completing this lesson, go to:

- Introduction to Confederations module

References

For additional information, refer to these resources:

- For more information on route reflectors, refer to “BGP Case Studies Section 4” at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#routereflectors
- For more information on configuring route reflectors, refer to “Configuring BGP” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2/1cfbgp.htm - xtocid45

Laboratory Exercise: BGP Route Reflectors

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will enable route reflectors to modify IBGP split-horizon rules in a transit AS.

After completing this exercise, you will be able to:

- Plan the migration of an existing IBGP backbone to a backbone with route reflectors
- Configure BGP route reflectors on a BGP backbone
- Monitor a BGP backbone containing route reflectors

 Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
bgp cluster-id *id*	Specify the cluster-ID.
neighbor *ip-address* **route-reflector-client**	Establish an IBGP session by using your workgroup number as the AS number.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- In this exercise, you will replace the IBGP full mesh with a hierarchy of BGP route reflectors.

Configure IBGP sessions in your AS according to this design:

- Cluster #1 is your top-level cluster. Within this cluster, router WGxR3 is the route reflector, and WGxR2 is the client.
- WGxR2 is the route reflector in cluster #2, and WGxR1 is its client.
- Router WGxR4 is not participating in any cluster and needs to be in IBGP full mesh with top-level route reflectors.
- The required IBGP sessions and the cluster structure are outlined in Figure 1.

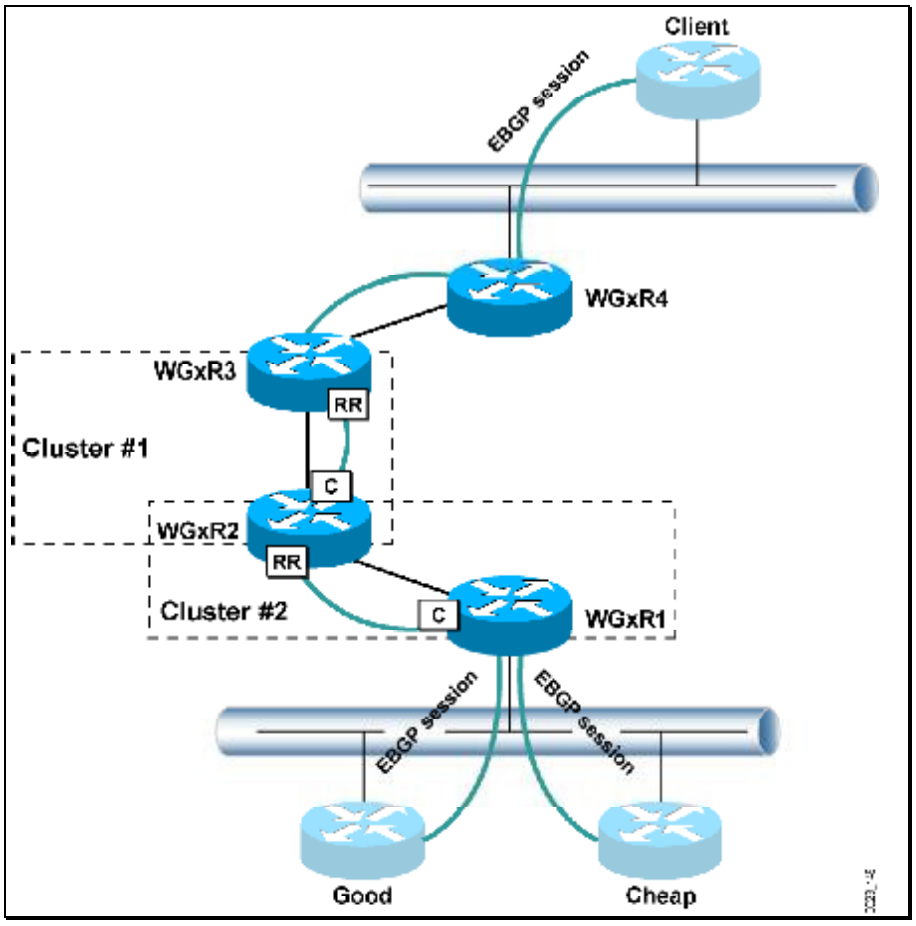


Figure 1: Implementing a BGP route reflector design

Exercise Procedure

Complete these steps:

Configure the top-level route reflector WGxR3:

Step 1 Configure cluster-ID 101 on the top-level route reflector WGxR3. Configure the client on WGxR3 and remove unnecessary IBGP sessions.

Configure the second-level route reflector WGxR2:

Step 2 Configure cluster-ID 102 on WGxR2. Configure the client on WGxR2 and remove unnecessary IBGP sessions from WGxR2.

Remove unnecessary BGP neighbors from your AS:

Step 3 Remove the IBGP sessions that are no longer needed from WGxR1 and WGxR4.

 Copyright © 2003, Cisco Systems, Inc.

Exercise Verification

You have completed this exercise when you attain these results:

- Check your BGP connections on the core routers and ensure that they are connected only to the peers, as shown in the lab diagram.

```
WG1R1#sh ip bgp summary
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.20.20   4    20    5923    4224      138    0    0 01:11:00
23
192.168.20.22   4    22    1720    1705      138    0    0 01:11:03
23
197.1.2.1       4     1    1665    1703      138    0    0 00:02:52
11


WG1R2#sh ip bgp sum
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
197.1.1.1       4     1    4149    4053     1074    0    0 00:03:19
25
197.1.4.1       4     1    4035    4053     1074    0    0 00:01:05
13


WG1R3#sh ip bgp sum
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
197.1.2.1       4     1    4053    4035     1028    0    0 00:01:26
25
197.1.6.1       4     1    4171    4068     1028    0    0 01:27:02
13


WG1R4#sh ip bgp sum
…
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.21.99   4    99    1744    1713      225    0    0 00:46:07
11
197.1.4.1       4     1    1674    1745      225    0    0 01:27:19
23
```

- Check to see if router “Client” is receiving only your networks and those originating in AS 37.

```
Client#sh ip bgp
BGP table version is 210, local router ID is 197.99.111.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 99.0.0.0         0.0.0.0                  0         32768 i
*> 128.37.0.0       192.168.21.1                           0 1 20 42
37 i
*> 192.37.11.0      192.168.21.1                           0 1 20 42
37 i
*> 192.168.1.0      192.168.21.1             0             0 1 i
*> 197.1.0.0/16     192.168.21.1             0             0 1 i
*> 197.99.1.0       0.0.0.0                  0         32768 i
*> 197.99.11.0      0.0.0.0                  0         32768 i
*> 197.99.12.0      0.0.0.0                  0         32768 i
*> 197.99.13.0      0.0.0.0                  0         32768 i
*> 197.99.20.0      0.0.0.0                  0         32768 i
*> 197.99.22.0      0.0.0.0                  0         32768 i
*> 197.99.80.0      0.0.0.0                  0         32768 i
*> 197.99.111.1/32  0.0.0.0                  0         32768 i
*> 197.99.120.0     0.0.0.0                  0         32768 i
*> 197.99.128.0/20  0.0.0.0                  0         32768 i
```

- Check to see if the cluster-list attribute is applied to routes passing through the AS.

```
WG1R1#sh ip bgp 197.99.1.0
BGP routing table entry for 197.99.1.0/24, version 129
Paths: (1 available, best #1, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  192.168.20.20 192.168.20.22
  99
    192.168.21.99 (metric 202) from 197.1.2.1 (197.1.7.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Originator: 197.1.7.1, Cluster list: 0.0.0.102, 0.0.0.101
```

Answer these questions:

Q1) Did this design require you to configure a cluster-ID?

Q2) What is the default cluster-ID?

Q3) When do you have to configure a cluster-ID?

 Copyright © 2003, Cisco Systems, Inc.

Introduction to Confederations

Overview

Routers within an autonomous system (AS) are typically configured in a full mesh. Confederations and route reflectors are Border Gateway Protocol (BGP) scalability mechanisms that enable routing information to be redistributed to all routers within an AS while eliminating the need for a fully meshed topology within the AS. These features reduce the number of TCP sessions that must be maintained, lowering network overhead, and CPU and memory requirements. Confederations can serve as an alternative or a complement to route reflectors. The confederation feature enables network administrators to break up an AS into a set of logical subautonomous systems.

This lesson introduces BGP confederations by explaining why confederations are used to improve BGP scalability. This lesson also discusses AS-path propagation and processing in an AS containing confederations.

Importance

Large BGP networks cannot properly scale without relying on performance-enhancing tools such as route reflectors and confederations. Confederations enable BGP routing information to be distributed in a fashion that does not require a physical full-mesh network. Network overhead is reduced by decreasing the number of TCP connections required to distribute routing information and by lessening router CPU and memory resource requirements.

Objectives

Upon completing this lesson, you will be able to:

- Explain the need for BGP confederations
- Explain the concept of AS splitting with BGP confederations
- Describe AS-path propagation in BGP confederations

- Explain AS-path attribute processing in an AS containing BGP confederations
- Explain the properties of intraconfederation EBGP sessions

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- BGP Transit Autonomous Systems module

Outline

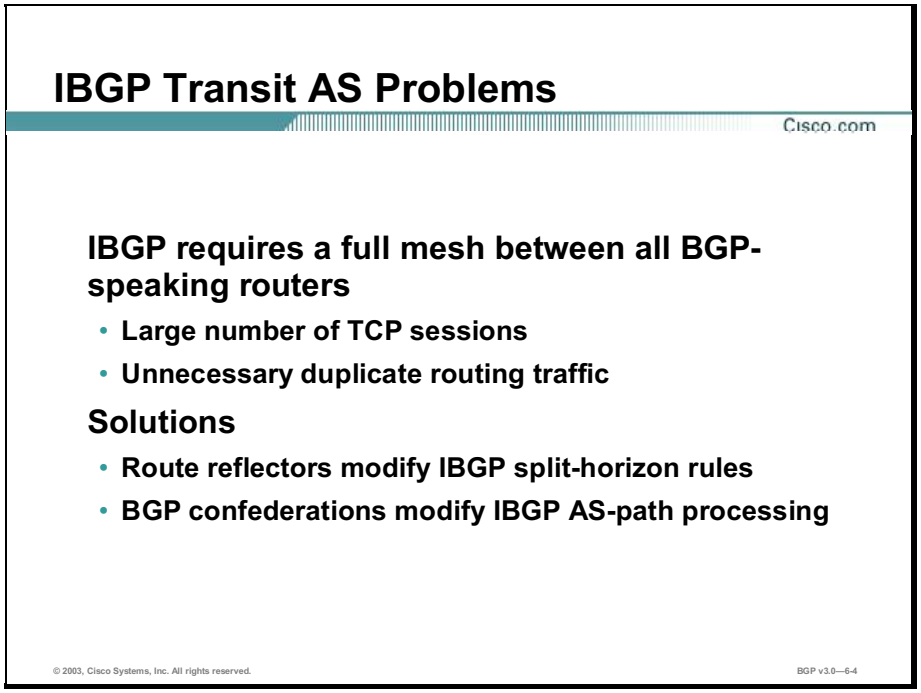
This lesson includes these topics:

- Overview
- IBGP Transit AS Problems
- Splitting a Transit AS with BGP Confederations
- AS-Path Propagation Within the BGP Confederation
- AS-Path Processing in BGP Confederations
- Intraconfederation EBGP Session Properties
- Summary
- Assessment (Quiz): Introduction to Confederations

 Copyright © 2003, Cisco Systems, Inc.

IBGP Transit AS Problems

This topic describes the Internal Border Gateway Protocol (IBGP) full-mesh requirement when you are using a transit AS, and the potential issues that this requirement can cause. Solutions to the IBGP full-mesh requirement are also discussed.


IBGP Transit AS Problems
Cisco.com
IBGP requires a full mesh between all BGP-speaking routers
• Large number of TCP sessions
• Unnecessary duplicate routing traffic
Solutions
• Route reflectors modify IBGP split-horizon rules
• BGP confederations modify IBGP AS-path processing
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-4

Classic IBGP split-horizon rules specify that BGP updates received on an External Border Gateway Protocol (EBGP) session should be forwarded on all IBGP and EBGP sessions, but BGP updates received on an IBGP session should be forwarded on all EBGP sessions only. This rule requires a boundary router to be able to update all other BGP-speaking routers in its own AS directly via an IBGP session established to each of them.

IBGP split-horizon rules avoid routing information loops within the AS. If IBGP information was forwarded to another IBGP peer router, the information might come back to the originator, and be forwarded again in a never-ending loop. The originator would not detect the loop because no BGP attributes are changed when an update is sent through an IBGP session.

The general design rule in classic IBGP is to have a full mesh of IBGP sessions between all BGP-speaking routers inside an AS. However, a full mesh of IBGP sessions between “n” number of routers would require (n * (n-1)) / 2 IBGP sessions. For example, an AS with 10 routers would require (10 * (10 – 1)) / 2 = 45 IBGP sessions.

Every BGP session between two routers is established through a separate TCP session to the BGP peer. An update that must be sent to all IBGP peers must be sent separately on each of the TCP sessions. If a router is attached to the rest of the network over just a single link, this single link has to carry all TCP/IP packets for all IBGP sessions. This requirement results in duplication of BGP updates over the single link.

Two different solutions are available to achieve greater scalability when you are faced with the full-mesh rules of IBGP autonomous systems:

- Route reflectors modify the classic IBGP split-horizon rules and allow a particular router to forward incoming IBGP updates to an outgoing IBGP session under certain conditions. This router becomes a concentration router, or a route reflector.

- BGP confederations introduce the concept of a number of smaller autonomous systems within the original AS. The small autonomous systems exchange BGP updates between them using intraconfederation EBGP sessions.

Example

A large service provider backbone, acting as a transit AS, contains 150 routers. Because the classic IBGP design rule requires a full mesh, the AS would require 11,175 sessions (n * [n-1] / 2). This number is impractical, and therefore you should use other solutions, such as BGP confederations, to reduce the full-mesh requirements of the network.

Practice

Q1) How do BGP confederations reduce the IBGP full mesh?

A) Confederations allow direct peering between external routers through a transit AS.

B) Confederations break an AS up into small “mini” autonomous systems forming EBGP sessions within the confederation.

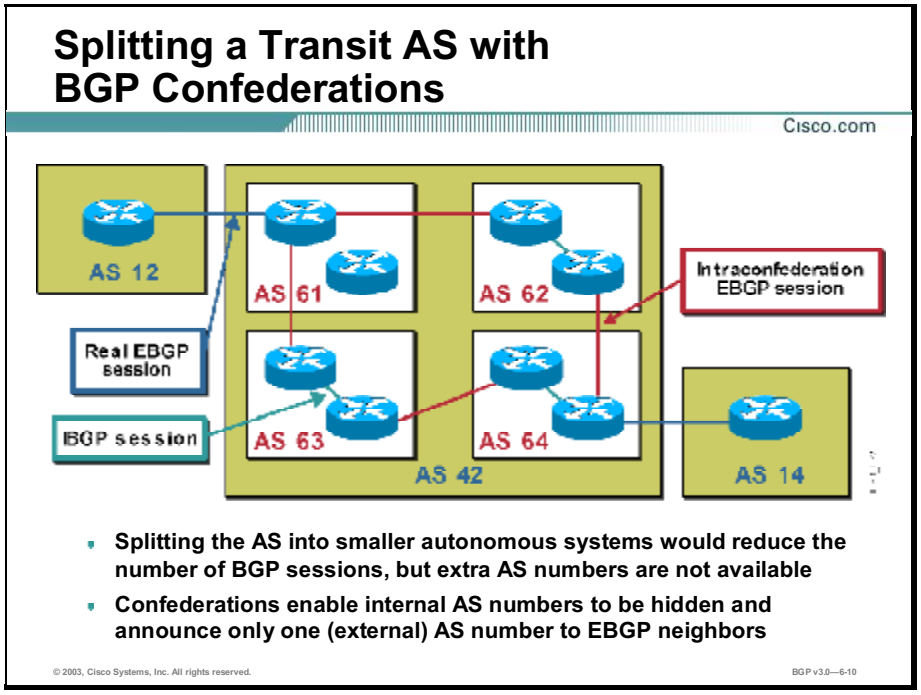
C) Confederations allow a router to advertise IBGP-learned routing information to other IBGP speakers, reducing the total number of peers required.

D) Confederations combine external autonomous systems into a larger “master” AS treating all internal routers as EBGP peers.

 Copyright © 2003, Cisco Systems, Inc.

Splitting a Transit AS with BGP Confederations

This topic describes how you can use BGP confederations to split an AS into a series of smaller autonomous systems.


Splitting a Transit AS with BGP Confederations
Cisco.com
AS 12
AS 61
AS 62
AS 63
AS 64
AS 42
AS 14
Intraconfederation EBGP session
Real EBGP session
IBGP session
Splitting the AS into smaller autonomous systems would reduce the number of BGP sessions, but extra AS numbers are not available
Confederations enable internal AS numbers to be hidden and announce only one (external) AS number to EBGP neighbors
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-10

A large number of routers in a large transit AS would traditionally introduce a complex full mesh structure of IBGP sessions. By splitting the AS into a number of small autonomous systems, you can provide each one of the small systems with a fairly simple IBGP structure. Interconnections between these autonomous systems could then be made using EBGP, which allows for arbitrary topologies.

Splitting an AS into smaller autonomous systems requires a large number of official AS numbers, which are a scarce resource.

By introducing the BGP confederation, a large AS is partitioned into a number of smaller autonomous systems (called “member autonomous systems”) where each is internal to the larger AS. The AS numbers of each member-AS used within the confederation are never visible from outside the confederation itself. This invisibility allows private AS numbers (in the range 64512-65535) to be assigned to autonomous systems inside a confederation to identify a member-AS, without the need to coordinate AS number assignments with an official AS delegation authority.

Within a member-AS, the classic IBGP rules apply. Therefore, all BGP routers inside the member-AS must still maintain a full mesh of BGP sessions.

Between member autonomous systems inside a confederation, EBGP sessions are established. These EBGP sessions behave slightly different from classic EBGP sessions and are therefore named intraconfederation EBGP sessions to differentiate them from true EBGP sessions.

Practice

Q1) What IBGP rule applies to member autonomous systems within a confederation?

A) The AS number in each member-AS must be from within the private AS number range.

B) Each member-AS within the confederation must use a globally registered AS number.

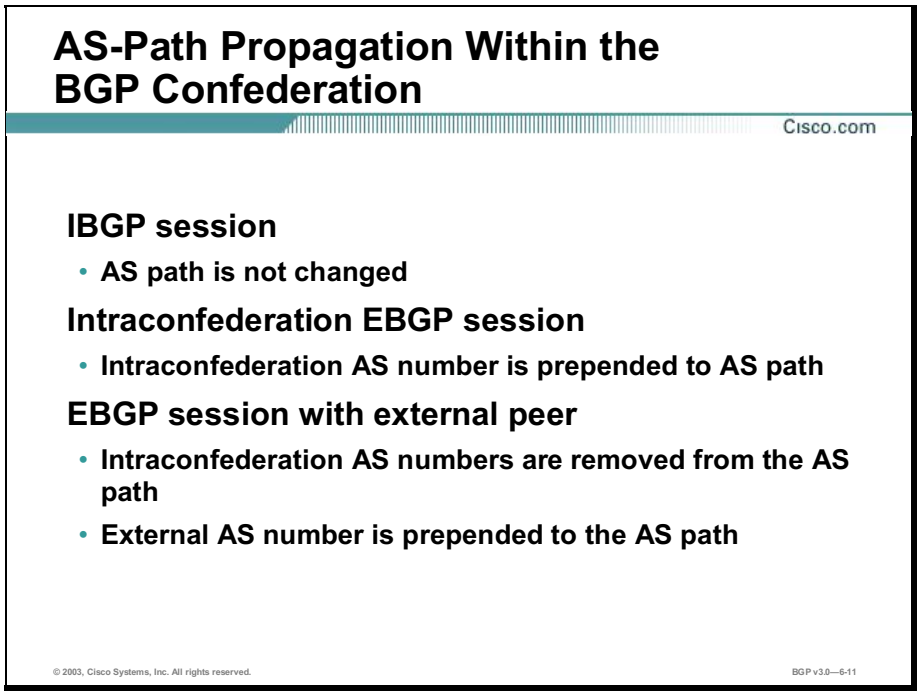
C) Each member-AS must be a transit AS for the confederation.

D) Routers contained within the member-AS must be fully meshed.

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Propagation Within the BGP Confederation

This topic describes how the AS-path attribute is propagated inside and outside of the BGP confederation.


AS-Path Propagation Within the BGP Confederation
Cisco.com
IBGP session
• AS path is not changed
Intraconfederation EBGP session
• Intraconfederation AS number is prepended to AS path
EBGP session with external peer
• Intraconfederation AS numbers are removed from the AS path
• External AS number is prepended to the AS path
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-11

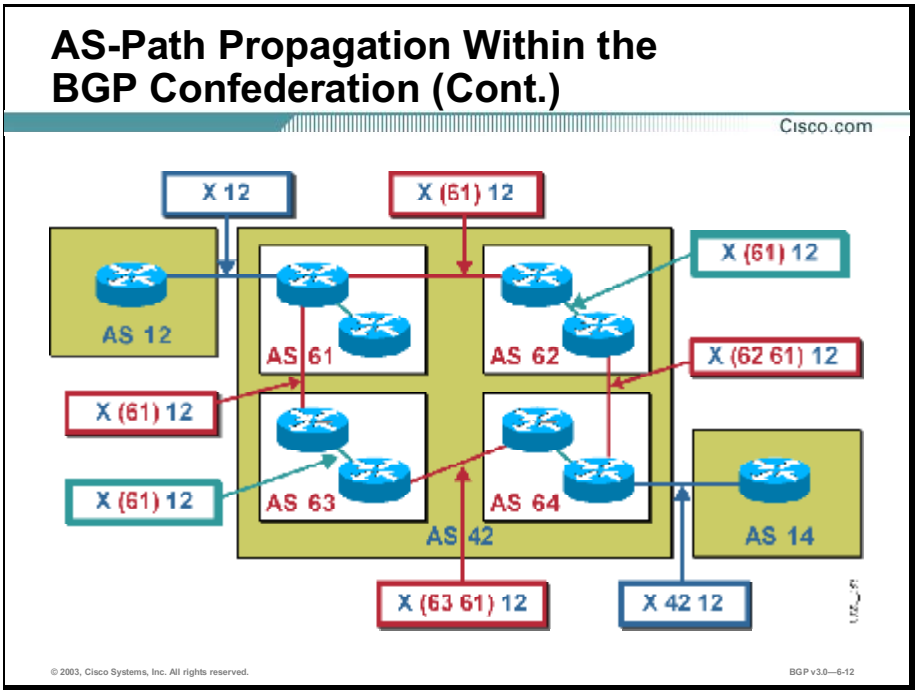
The mandatory well-known BGP attribute AS-path is modified on EBGP and intraconfederation EBGP sessions. The sender prepends its own AS number to the AS path whenever an EBGP update is sent. When a BGP update traverses the Internet, each and every AS it passes through will be recorded in the AS path. If the update, for any reason, comes back to an AS in which it has already been, the receiving router will recognize its own AS in the AS path and silently ignore the update. This mechanism prevents information loops and allows arbitrary topology when interconnecting autonomous systems.

IBGP sessions do not modify the AS-path attribute, so the topology within each AS is limited to the full mesh, and the propagation of BGP updates across multiple IBGP sessions is prohibited.

When a router sends a BGP update over an intraconfederation EBGP session, it prepends the member-AS number to the AS path. This information is maintained by the routers within the confederation and prevents routing information loops inside the confederation.

When a router sends a BGP update over a true EBGP session to an AS outside of the confederation, it removes the part of the AS path describing the member-AS numbers and prepends the official AS number to the AS path. As a result, the confederation appears as one single AS to the outside world.

Example



The figure illustrates how the AS-path attribute is processed within a BGP confederation.

Network X originates inside AS 12 and is announced by the edge router in AS 12 over a true EBGP session to the ingress router in the confederation. The edge router in AS 12 determines that the edge router it is communicating with resides in AS 42. AS 12 prepends its assigned AS number in the AS path (which was previously empty, because network X originated in AS 12). When the EBGP update arrives at the confederation member-AS 61, the AS-path attribute has been set to 12.

The member-AS 61 has intraconfederation EBGP sessions to member-AS 62 and to member-AS 63. The router in AS 61 prepends its own AS number to the AS path. When doing this, it signals that this part of the AS path describes the intraconfederation AS path. When printed out, the intraconfederation part of the AS path is displayed within parentheses. Therefore, when member AS 61 sends the update to member-AS 62 and member-AS 63, the route to network X has an AS-path attribute set to (61) 12.

Within a member-AS, the router sends the update using classic IBGP. The router does not modify the AS path when transmitting it over an IBGP session.

Member-AS 62 and member-AS 63 both prepend their AS number, so member-AS 64 receives the update about the route to X via two different paths, one with AS path (62 61) 12 and the other with AS path (63 61) 12.

Member-AS 64 selects one of the alternatives as the best BGP route. It then forwards this update on the intraconfederation EBGP session. This update could introduce a loop, but if the update was ever to be forwarded all the way back up to member-AS 61, the loop would be detected and member-AS 61 would silently ignore the update.

 Copyright © 2003, Cisco Systems, Inc.

Member-AS 64 also forwards the update about network X on a true EBGP session to AS 14. When it does, it removes all the parenthesized information in the AS path and replaces it with the official AS number of the confederation, 42. AS 14 thus receives the update about network X with an AS path set to 42 12.

The routers in AS 14 select the best path based on the length of the AS path if no other policy is configured. AS14 will see the route to X with an AS-path length of two. When AS 14 forwards packets destined for network X into the confederation (AS 42), member-AS 64 must make a forwarding decision, which is consistent with that of AS 14, to avoid a routing loop. Therefore, when a router within the BGP confederation AS 64,compares the AS-path length of two alternate paths, it does not account for the intraconfederation part of the AS- path attribute.

Practice

Q1) How is the AS path propagated between intraconfederation EBGP peers?

A) The confederation AS number is prepended to the AS-path attribute to ensure that the AS path is valid when the route update is sent to the EBGP peer in another AS.

B) Intraconfederation EBGP peers maintain a second AS path so that IBGP mesh requirements can be maintained without affecting the AS path used to reach EBGP peers.

C) The member-AS number is prepended to the AS-path attribute.

D) Because the confederation resides in a single AS, IBGP rules apply, and no AS number is prepended to the AS-path attribute.

Q2) What happens to the AS path when a BGP update is sent to an external peer outside of the confederation?

A) The AS path is not modified because it already contains the AS numbers of each member-AS.

B) The AS number of the confederation is prepended to the existing AS path.

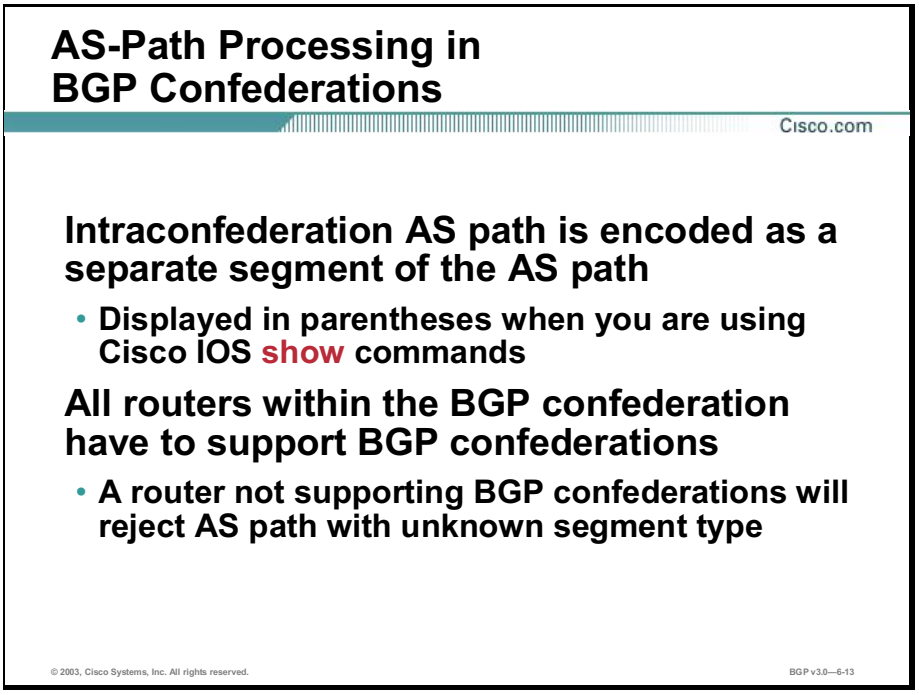
C) The intraconfederation AS numbers are removed from the path, and the external AS is prepended.

D) All member-AS entries are removed from the AS-path attribute by the EBGP peer.

Copyright © 2003, Cisco Systems, Inc.

AS-Path Processing in BGP Confederations

This topic explains how the BGP AS-path attribute is processed within an AS containing a BGP confederation.


AS-Path Processing in BGP Confederations
Cisco.com
Intraconfederation AS path is encoded as a separate segment of the AS path
• Displayed in parentheses when you are using Cisco IOS show commands
All routers within the BGP confederation have to support BGP confederations
• A router not supporting BGP confederations will reject AS path with unknown segment type
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-13

When BGP routing updates are sent by BGP speaking routers over a BGP session, the BGP attributes are encoded in a binary structure. The AS-path attribute, which is printed out and displayed as a text string, is actually a type, length, value (TLV) binary field, composed of several segments. The intraconfederation part of the AS path is encoded by the intraconfederation router as a separate segment of the AS path with a new type code. This segment of the AS path contains a sequence of AS numbers encoding the member autonomous systems that the BGP update has traversed.

Because this segment is an extension to the original interpretation of the mandatory well-known BGP attribute AS-path, a BGP implementation that does not support BGP confederations will not understand the intraconfederation part of the AS path. If a router receives a BGP update with a mandatory well-known attribute that the router cannot interpret, it will send a notification to the neighbor sending the offending update and terminate the session. A router not supporting BGP confederations, therefore, cannot operate inside a BGP confederation.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What happens if a router not supporting BGP confederations is placed inside the confederation?

A) The router will learn about the confederation through BGP capabilities exchange.

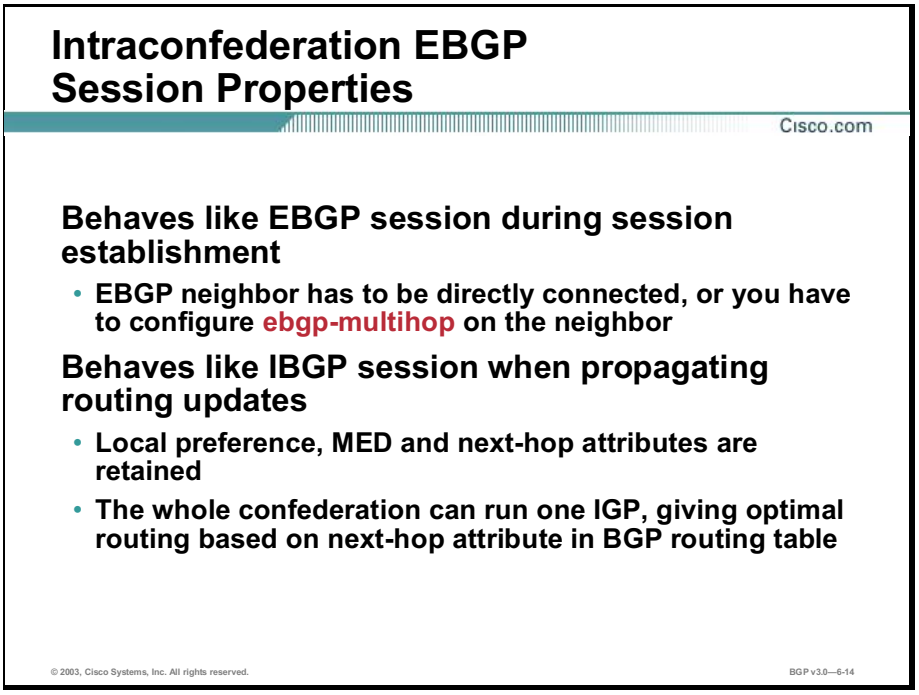
B) The router will process BGP updates as normal because it has to be aware only of the member-AS to which it belongs.

C) The router will automatically convert the intraconfederation AS numbers to the external AS number of the confederation.

D) The router will be unable to interpret the AS-path attribute and terminate its BGP session with that peer.

Intraconfederation EBGP Session Properties

This topic describes EBGP sessions between different autonomous systems contained within the confederation.


Intraconfederation EBGP Session Properties
Cisco.com
Behaves like EBGP session during session establishment
• EBGP neighbor has to be directly connected, or you have to configure ebgp-multihop on the neighbor
Behaves like IBGP session when propagating routing updates
• Local preference, MED and next-hop attributes are retained
• The whole confederation can run one IGP, giving optimal routing based on next-hop attribute in BGP routing table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-14

Intraconfederation EBGP sessions, while having EBGP-like properties (for example, updating the AS-path attribute when propagating BGP routes), still run inside a real AS and therefore have to share some properties with IBGP sessions to achieve the same end results. Similar to IBGP sessions, the BGP attributes of local preference, multi-exit discriminator (MED), and next-hop are not changed in updates propagated across intraconfederation EBGP sessions. All routers in all member autonomous systems inside the confederation consequently use the same next-hop address when they are doing recursive routing. Since all intraconfederation routers use the same next-hop address, the entire confederation should use the same IGP to resolve the BGP next-hop address. The IGP information should not be limited by the member-AS boundary.

On the other hand, intraconfederation EBGP sessions behave exactly like EBGP sessions when they are established. EBGP sessions are normally opened between directly connected interfaces. However, because all routers within the confederation run the same IGP and exchange internal routing information, there is no problem for them to open multihop sessions. Resilience of BGP sessions and consequent stability of BGP routing are introduced into the network if the intraconfederation EBGP sessions are established between loopback interfaces, just like IBGP sessions normally are.

When intraconfederation EBGP sessions are opened between loopback interfaces, the **ebgp-multihop** qualifier must be given to the session. Otherwise the EBGP session will never leave the Idle state.

 Copyright © 2003, Cisco Systems, Inc.

Actually, the intraconfederation EBGP sessions could be established between intraconfederation routers in an arbitrary topology, not necessarily following the physical topology. The next hop of the route will always contain the IP address of a BGP router outside of the confederation, and packet forwarding will follow the optimal path, because recursive routing will rely on the IGP to reach the BGP next hop. The intraconfederation EBGP sessions are merely used to distribute the BGP updates to all member autonomous systems. To avoid unnecessary duplication of routing updates, network designers should take great care when designing the topology of the intraconfederation EBGP sessions.

Practice

Q1) What is the difference between an EBGP session and intraconfederation EBGP session?

A) EBGP and intraconfederation EBGP sessions are identical in how they treat EBGP peers.

B) EBGP sessions do not allow sessions to peers that are not directly attached.

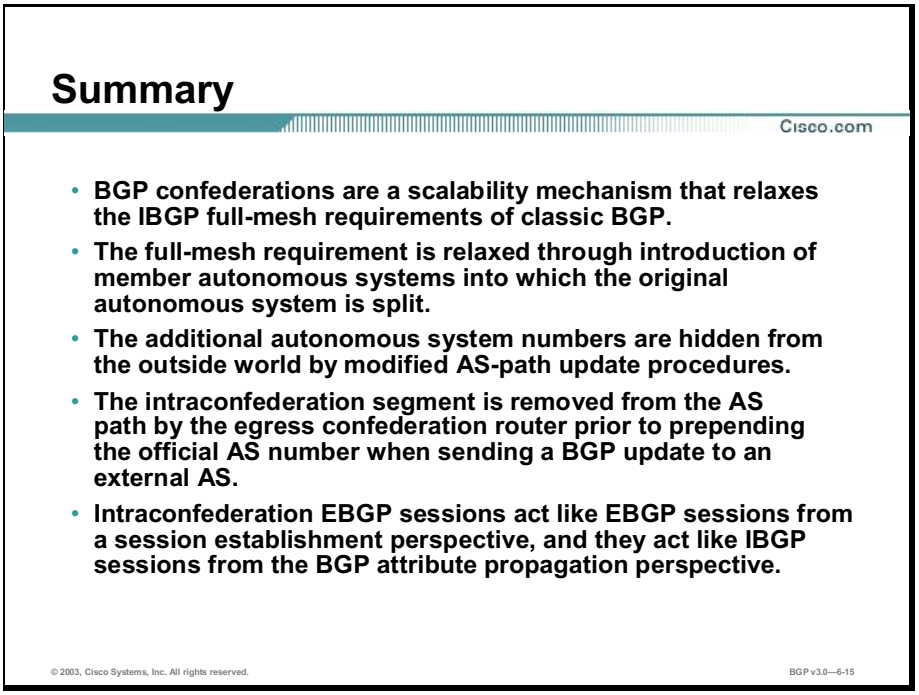
C) Intraconfederation EBGP routers retain IBGP processing rules of the local preference, MED, and next-hop attributes.

D) Intraconfederation EBGP routers can form BGP peering relationships only between member-AS routers.

Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• BGP confederations are a scalability mechanism that relaxes the IBGP full-mesh requirements of classic BGP.
• The full-mesh requirement is relaxed through introduction of member autonomous systems into which the original autonomous system is split.
• The additional autonomous system numbers are hidden from the outside world by modified AS-path update procedures.
• The intraconfederation segment is removed from the AS path by the egress confederation router prior to prepending the official AS number when sending a BGP update to an external AS.
• Intraconfederation EBGP sessions act like EBGP sessions from a session establishment perspective, and they act like IBGP sessions from the BGP attribute propagation perspective.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-15

Next Steps

After completing this lesson, go to:

- Configuring and Monitoring Confederations lesson

References

For additional information, refer to these resources:

- For more information on BGP confederations, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

 Copyright © 2003, Cisco Systems, Inc.

Quiz: Introduction to Confederations

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Explain the need for BGP confederations
- Explain the concept of AS splitting with BGP confederations
- Describe AS-path propagation in BGP confederations
- Explain AS-path attribute processing in an AS containing BGP confederations
- Explain the properties of intraconfederation EBGP sessions

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What is the main problem solved by implementing BGP confederations?

A) the large number of routes carried in the IGP when BGP is deployed

B) the ability for BGP to scale a single AS in a large network

C) the need for a homogeneous method of applying policies to routes carried through an AS

D) the lack of a service level agreement supporting features and the need to deploy these features with greater ease

Copyright © 2003, Cisco Systems, Inc.

Q2) Although confederations eliminate the need for a fully meshed topology within the AS, where does the BGP full-mesh requirement still apply?

A) To all EBGP neighbor sessions.

B) Inside each member-AS.

C) Between the member autonomous systems contained in the confederation.

D) The IBGP full-mesh requirement no longer applies when confederations are used in an AS.

Q3) How does an IBGP router receiving the AS-path attribute in a BGP update determine if the route has crossed a member-AS within a confederation?

A) By the presence of the confederation bit in the flag field of the BGP update.

B) Because the AS-path attribute will contain only the AS number of the ingress EBGP peer.

C) The member-AS numbers will be indicated by the presence of parentheses surrounding the AS number entry.

D) The IBGP router cannot determine whether the route has crossed a member-AS, because the AS number of each AS boundary crossed is appended to the AS-path attribute.

Q4) How does an EBGP router receiving the AS-path attribute in a BGP update determine if the route has crossed a member-AS within a confederation?

A) By the presence of the confederation bit in the flag field of the BGP update.

B) Because the AS-path attribute will contain only the AS number of the ingress EBGP peer.

C) The member-AS numbers will be indicated by the presence of parentheses surrounding the AS number entry.

D) The EBGP router cannot determine whether the route has crossed a member-AS, because the member-AS entries are removed from the AS path prior to exiting the confederation.

 Copyright © 2003, Cisco Systems, Inc.

Q5) Why is it not possible for a router that does not support BGP confederations to operate inside an AS configured as a confederation?

A) The router will believe that the AS path is longer than the actual AS path and route incorrectly.

B) The router will be unable to interpret the AS-path attribute and terminate its BGP session with that peer.

C) The router will automatically convert the intraconfederation AS numbers to the external AS number of the confederation, causing an AS number mismatch.

D) The router will process BGP updates as normal because it has to be aware only of the member-AS to which it belongs, causing incorrect routing information to propagate through the AS.

Q6) Which three IBGP properties are retained within the confederation even though EBGP sessions between member autonomous systems are formed? (Choose three.)

A) local preference

B) MED

C) weight

D) next-hop

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Configuring and Monitoring Confederations

Overview

This lesson introduces the steps required to successfully migrate an existing autonomous system (AS) to Border Gateway Protocol (BGP) confederations. It also presents the Cisco IOS® commands required to configure and monitor confederations.

Importance

Large BGP networks cannot properly scale without relying on performance-enhancing tools such as route reflectors and confederations. Confederations enable BGP routing information to be distributed in a fashion that does not require a physical full-mesh network. Implementing such a network requires knowledge of the steps to properly migrate and configure BGP confederations, and the commands used to verify the operation of the configured network.

Objectives

Upon completing this lesson, you will be able to:

- List the design rules for migrating an existing IBGP backbone to BGP confederations
- Plan the migration of an existing IBGP backbone toward BGP confederation
- Identify the Cisco IOS commands required to configure BGP confederations on a BGP backbone
- Identify the Cisco IOS commands required to monitor a BGP backbone containing BGP confederations

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- BGP Transit Autonomous Systems module

Outline

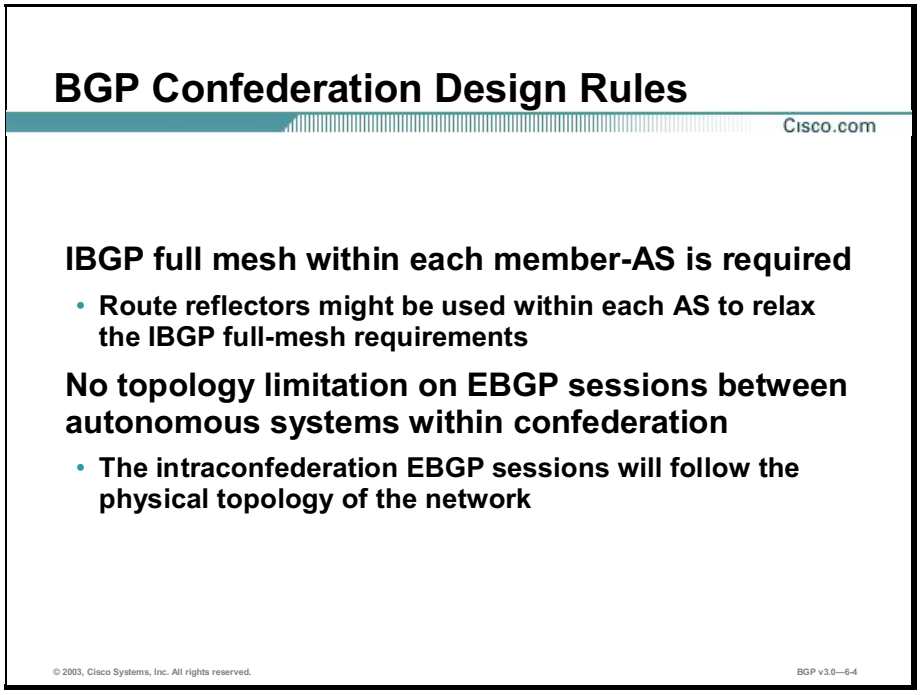
This lesson includes these topics:

- Overview
- BGP Confederation Design Rules
- Planning BGP Confederations
- Configuring BGP Confederations
- Monitoring BGP Confederations
- Summary
- Assessment (Lab): BGP Confederations

 Copyright © 2003, Cisco Systems, Inc.

BGP Confederation Design Rules

This topic describes the basic design rules that network designers should follow when planning a transit AS for BGP confederations.


BGP Confederation Design Rules
Cisco.com
IBGP full mesh within each member-AS is required
• Route reflectors might be used within each AS to relax the IBGP full-mesh requirements
No topology limitation on EBGP sessions between autonomous systems within confederation
• The intraconfederation EBGP sessions will follow the physical topology of the network
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-4

When designing BGP confederations, keep in mind two basic design rules:

- There are no restrictions on intraconfederation External Border Gateway Protocol (EBGP) sessions.
- A full mesh of Internal Border Gateway Protocol (IBGP) sessions is still needed inside every member-AS.

BGP confederations do not modify IBGP behavior and therefore the classic IBGP split-horizon rules still apply. As a result, a full mesh of IBGP sessions between all routers in the member-AS is required. The basic idea of BGP confederations is to make the member-AS smaller than the original larger AS so that the full mesh will not be too complex. Route reflector functionality is also available within a member-AS to reduce the complexity of IBGP sessions if needed.

In theory, the member autonomous systems may be interconnected by intraconfederation EBGP sessions in an arbitrary topology. However, a structure that is too complicated introduces unnecessary duplication of information. There is no split-horizon function on EBGP sessions, meaning that in a redundant configuration, where a BGP update may be forwarded by a router over two different member autonomous systems to reach a third member-AS, the receiving member-AS will get both copies of the update.

Experience shows that a centralized confederation design leads to the best behavior. Centralized design means that all member autonomous systems will exchange information with each other via a central member-AS backbone.

On the other hand, too few intraconfederation EBGP sessions may introduce single points of failure. If two member autonomous systems are redundantly connected on the physical level but have only a single intraconfederation EBGP session between them, a single point of failure is introduced.

Good network design should never introduce additional single points of failure. If the physical topology is redundant, then the intraconfederation EBGP sessions should be redundant as well.

In a hierarchical network topology, the network core could serve as a central member-AS backbone. The more peripheral parts of the network could be divided into several member autonomous systems that are all connected to the central member-AS.

Practice

Q1) How can you reduce the IBGP full mesh within a confederation AS?

A) You cannot reduce the full mesh because all IBGP peers must be fully meshed within the member-AS.

B) Implementing router reflectors inside a member-AS can reduce the IBGP full-mesh requirement.

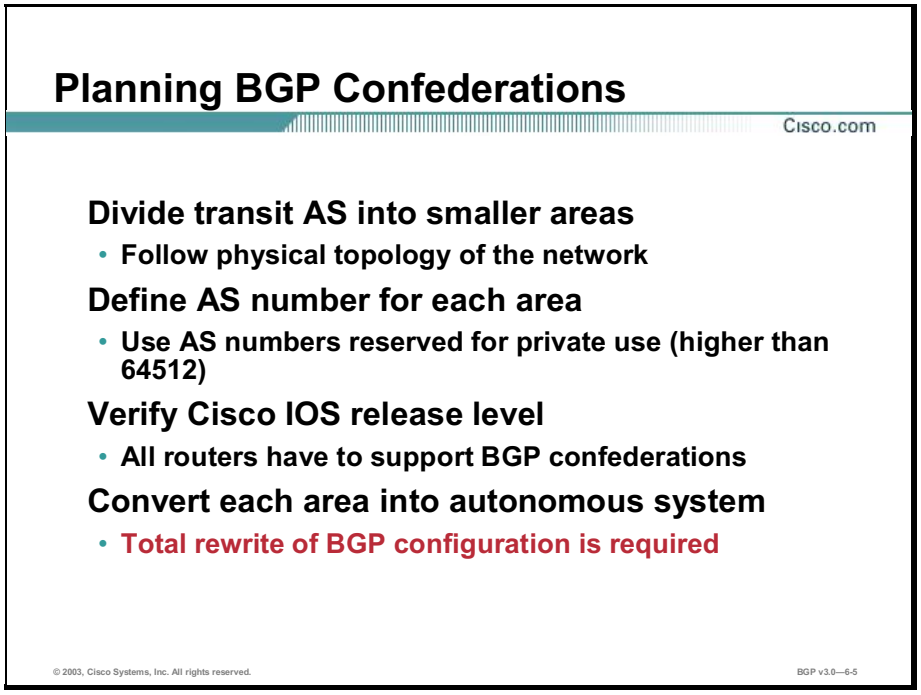
C) You can nest a confederation within a confederation to remove the IBGP full-mesh requirement.

D) Because confederations are used, there is no requirement for an IBGP full mesh within each member-AS.

 Copyright © 2003, Cisco Systems, Inc.

Planning BGP Confederations

This topic describes how to plan a BGP backbone for a configuration that includes BGP confederations.


Planning BGP Confederations
Cisco.com
Divide transit AS into smaller areas
• Follow physical topology of the network
Define AS number for each area
• Use AS numbers reserved for private use (higher than 64512)
Verify Cisco IOS release level
• All routers have to support BGP confederations
Convert each area into autonomous system
• Total rewrite of BGP configuration is required
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-5

The physical topology of the AS serves as a guide to design the confederation. You should introduce no additional single points of failure when implementing the confederation. If the physical topology is redundant, good practice is to have redundant intraconfederation EBGP sessions. If the physical topology is not redundant, introducing a nonredundant set of sessions does not add a single point of failure—the network was already nonredundant.

You need to make the following preparations before migrating from a full mesh of IBGP sessions to a confederation design:

- Identify a group of core routers that can serve as a central member-AS.
- Identify several groups of more peripheral routers where within each group, routers are well connected. Let each group be its own member-AS.
- Make a plan how to assign AS numbers (64512 – 65535) to your member-AS. The plan must uniquely identify each member-AS within the confederation. Each member-AS is not seen from outside the confederation, so you do not need to coordinate the plan with any other AS.
- Make sure that no router is lacking support (correct Cisco IOS release level) for a BGP confederation. If any router does, it will break the network.
- Remove the original BGP configuration with the original official AS number.

Note	Support for the BGP confederation feature is included in Cisco IOS software since version 10.3.

Practice

Q1) Which three requirements are BGP confederation planning steps? (Choose three.)

A) Define an AS number for each member-AS.

B) Remove the IBGP mesh in each member-AS.

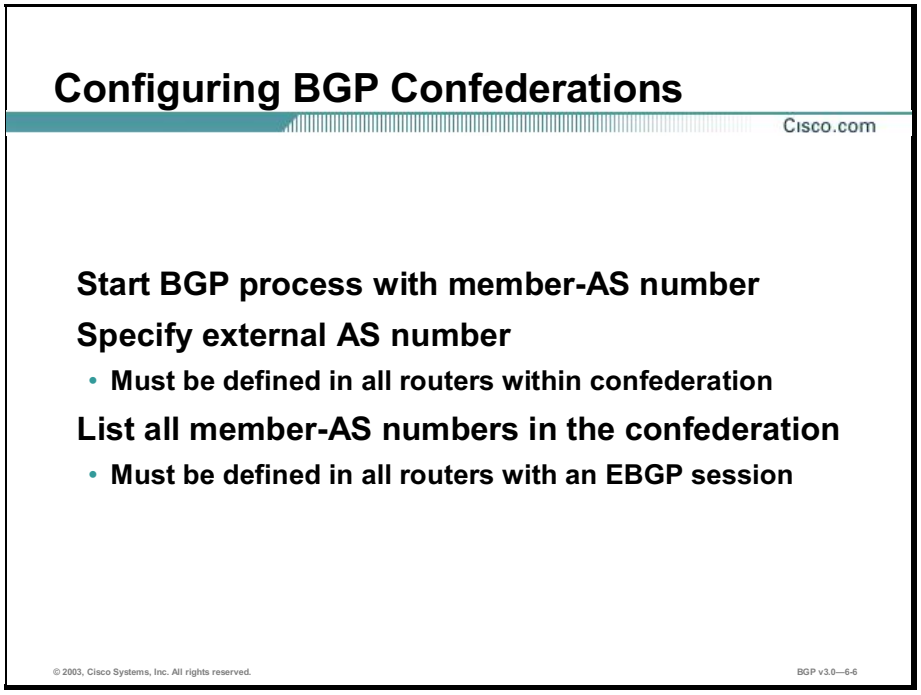
C) Divide the transit AS into smaller autonomous systems.

D) Verify that the Cisco IOS version installed on the router supports confederations.

 Copyright © 2003, Cisco Systems, Inc.

Configuring BGP Confederations

This topic describes how to configure BGP confederations on a BGP backbone. An example of a BGP confederations configuration is also included in this topic.

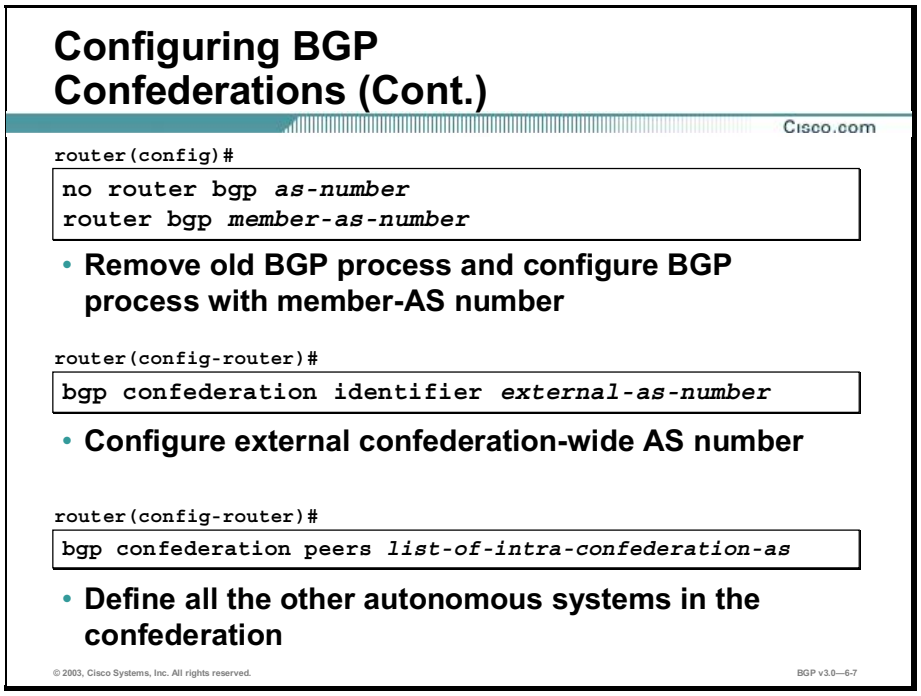

Configuring BGP Confederations
Cisco.com
Start BGP process with member-AS number
Specify external AS number
• Must be defined in all routers within confederation
List all member-AS numbers in the confederation
• Must be defined in all routers with an EBGP session
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-6

After you have done the preparation necessary to migrate from a full mesh of IBGP sessions to a confederation design, you need to complete the following configuration steps:

- Make a new BGP configuration using the internal member-AS number according to the AS number plan for the confederation.
- Specify the original official AS number as the identifier of the confederation. This information will be used by egress confederation routers whenever communicating with other external autonomous systems.
- Specify a list of the member-AS numbers being used. The router uses this information to distinguish between intraconfederation EBGP behavior and true EBGP behavior.
- Configure all the IBGP sessions in the full mesh within the member-AS.
- Configure intraconfederation EBGP sessions between each member-AS in a way that introduces no additional single point of failure.
- Configure true EBGP sessions with external autonomous systems.

Note	Removing the original BGP configuration and creating a new BGP configuration will always cause interruption in network availability. Migration to BGP confederation has to be a well-planned process.

 Copyright © 2003, Cisco Systems, Inc.


Configuring BGP Confederations (Cont.)
Cisco.com
router(config)#
no router bgp as-number
router bgp member-as-number
• Remove old BGP process and configure BGP process with member-AS number
router(config-router)#
bgp confederation identifier external-as-number
• Configure external confederation-wide AS number
router(config-router)#
bgp confederation peers list-of-intra-confederation-as
• Define all the other autonomous systems in the confederation
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-7

router bgp

To configure the BGP routing process, use the **router bgp** global configuration command.

router bgp *as-number*

To remove a routing process, use the **no** form of this command.

no router bgp *as-number*

Syntax Description

as-number	Number of an AS that identifies the router to other BGP routers and tags the routing information passed along

Note The AS number specified when you are configuring a BGP process inside a confederation is the intraconfederation (member) AS number.

bgp confederation identifier

To specify a BGP confederation identifier, use the **bgp confederation identifier** router configuration command.

bgp confederation identifier *external-as-number*

To remove the confederation identifier, use the **no** form of this command.

no bgp confederation identifier *external-as-number*

Syntax Description

external-as-number	Public AS number that the confederation is using externally

bgp confederation peers

To configure the autonomous systems that belong to the confederation, use the **bgp confederation peers** router configuration command.

bgp confederation peers *as-number* [*as-number*]

To remove an AS from the confederation, use the **no** form of this command.

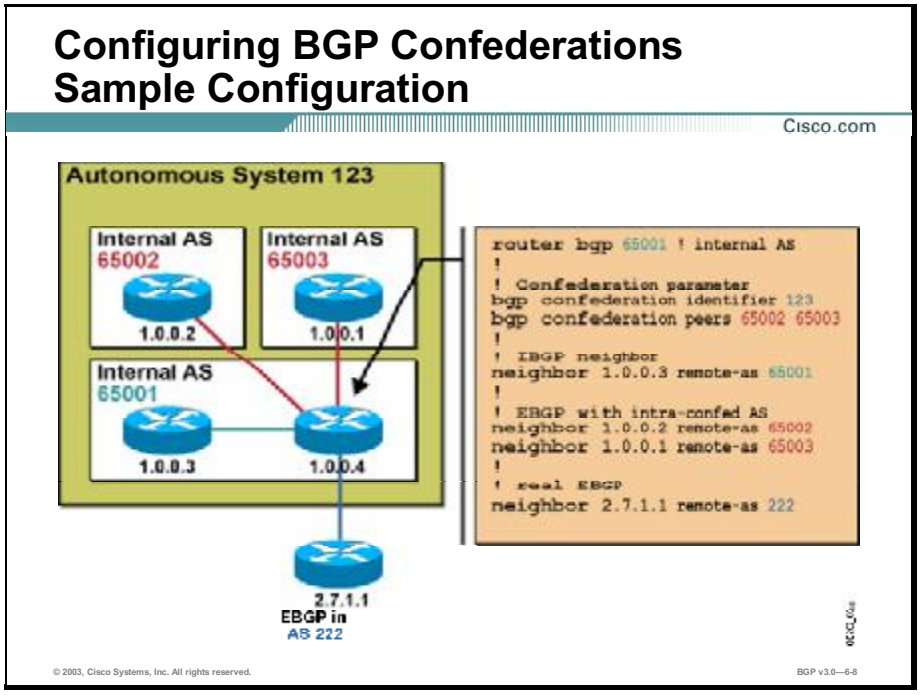
no bgp confederation peers *as-number* [*as-number*]

Syntax Description

as-number	AS numbers of member autonomous systems inside the confederation. This list is used by the router to distinguish an intraconfederation EBGP session from a real EBGP session with an external AS.

 Copyright © 2003, Cisco Systems, Inc.

Example



The example here illustrates the use of confederation-related BGP configuration commands.

Note	The example shows only a portion of the configuration of router 1.0.0.4. The configuration displayed is not a complete configuration—only those parts that are relevant to BGP confederations are displayed.

AS 123 is transformed into a BGP confederation. The member-AS 65001 serves as a central member-AS. Member autonomous systems 65002 and 65003 are connected to AS 65001. From the outside, the confederation looks like a single AS, still identified by AS number 123.

The internal member-AS number is specified in the **router bgp** configuration command. This number is the AS number that will be prepended to the AS-path attribute by the router when updates are forwarded across intraconfederation EBGP sessions.

The original, official AS number is given as the BGP confederation identifier. This number is the AS number that will replace the internal member-AS information when egress confederation routers forward updates across EBGP sessions with external autonomous systems.

The **bgp confederation peers** configuration step allows the router to identify the type of session to be opened with the BGP peers configured through the **neighbor** statement.

Neighbor 1.0.0.3 is in the same member-AS, so it is an IBGP session. Both neighbor 1.0.0.2 (AS 65002) and 1.0.0.1 (AS 65003) belong to different member-autonomous systems (the AS number appears in the list of BGP confederation peers), so they are intraconfederation EBGP sessions.

Neighbor 2.7.1.1 belongs to AS 222, which is not listed as a member-AS in the **bgp confederation peers** command, so this is a true EBGP session.

Practice

Q1) Which two BGP parameters do you need to specify on every router within a confederation? (Choose two.)

A) a list of all AS numbers in the confederation

B) a list of all true EBGP sessions

C) the official AS number (as the identifier of the confederation)

D) the correct MD5 authentication password in each peer

Q2) What impact will configuring BGP confederations have on the BGP network?

A) BGP update performance will be increased because the member-AS mesh is reduced.

B) EBGP peers will have a reduced number of **neighbor** statements to connect to the transit backbone.

C) Network availability will be impacted because the BGP routing process must be removed and restored using the new AS numbers.

D) IBGP session establishment times will be reduced.

Q3) What does the BGP confederation identifier define?

A) the AS number of the confederation external peer

B) the public AS number that the confederation is using externally

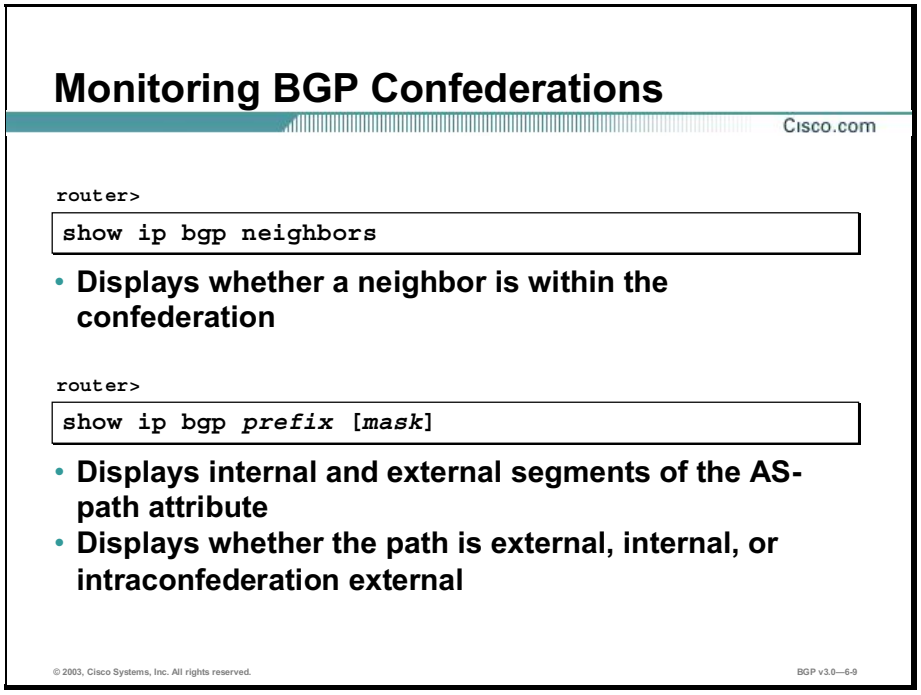
C) the AS number of the member-AS

D) the MD5 authentication password for the confederation

 Copyright © 2003, Cisco Systems, Inc.

Monitoring BGP Confederations

This topic explains the basic commands used to monitor a BGP backbone that you have configured with BGP confederations.


Monitoring BGP Confederations
Cisco.com
router>
show ip bgp neighbors
• Displays whether a neighbor is within the confederation
router>
show ip bgp prefix [mask]
• Displays internal and external segments of the AS-path attribute
• Displays whether the path is external, internal, or intraconfederation external
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-9

show ip bgp neighbors

To display information about the TCP and BGP connections to neighbors, use the **show ip bgp neighbors** EXEC command.

show ip bgp neighbors [*address*] [**received-routes** | **routes** | **advertised-routes** | {**paths** *regular-expression*} | **dampened-routes**]

Syntax Description

address	Address of a specific neighbor you wish to display information about. If you omit this argument, all neighbors are displayed.
received-routes	Displays all received routes (both accepted and rejected) from the specified neighbor.
routes	Displays all routes that are received and accepted. This is a subset of the output from the **received-routes** keyword.
advertised-routes	Displays all the routes that the router has advertised to the neighbor.
paths regexp	Regular expression that is used to match the paths received.
dampened-routes	Displays the dampened routes to the neighbor at the IP address specified.

If the **show ip bgp neighbors** command is executed on a router without any keywords, the resulting information that is displayed does not show routes or paths received by the router, but instead shows the status of its neighbor sessions.

show ip bgp

To display entries in the BGP routing table, use the **show ip bgp** EXEC command.

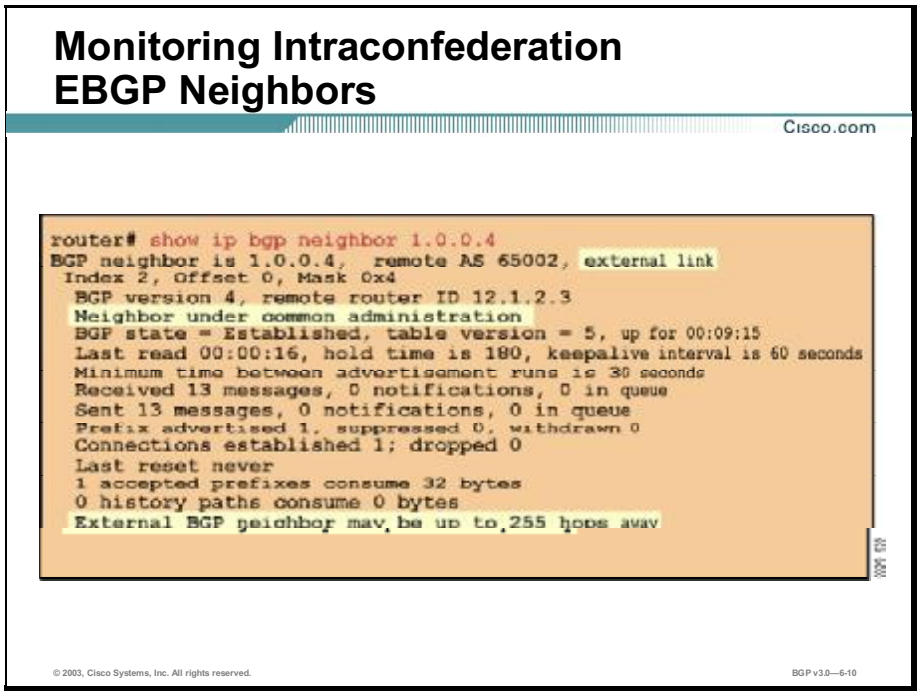
show ip bgp [*network*] [*network-mask*] [**longer-prefixes**]

When details are displayed for a specific route entry in the BGP table, the next hop and AS path are displayed along with information indicating whether a BGP update was received over an intraconfederation EBGP session or a regular EBGP session.

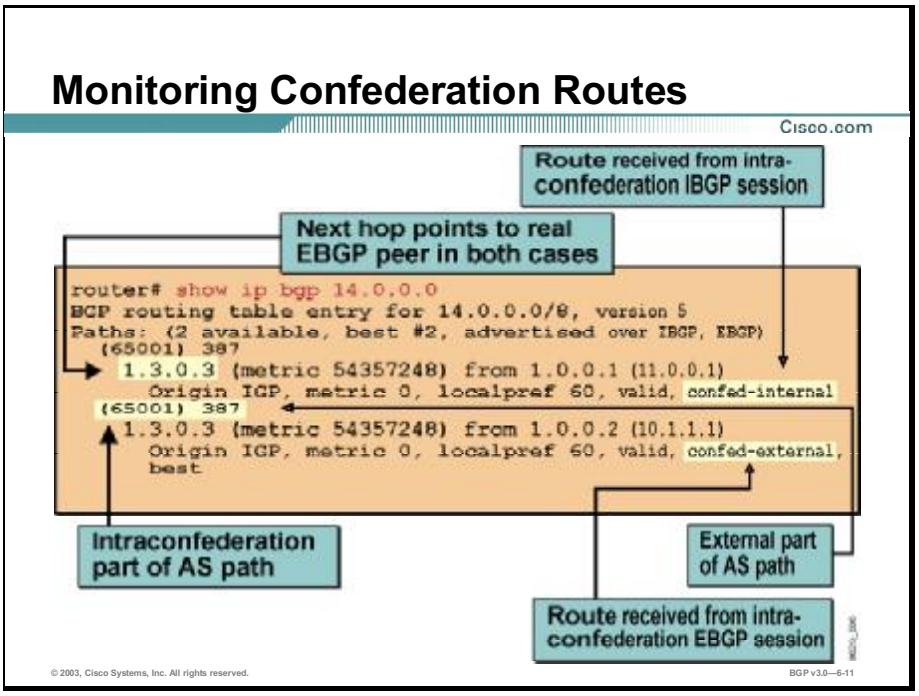
Syntax Description

network	Network number, entered to display a particular network in the BGP routing table
network-mask	Displays all BGP routes matching the address and mask pair
longer-prefixes	Displays the route and more specific routes

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Intraconfederation EBGP Neighbors
Cisco.com
router# show ip bgp neighbor 1.0.0.4
BGP neighbor is 1.0.0.4, remote AS 65002, external link
Index 2, Offset 0, Mask 0x4
BGP version 4, remote router ID 12.1.2.3
Neighbor under common administration
BGP state = Established, table version = 5, up for 00:09:15
Last read 00:00:16, hold time is 180, keepalive interval is 60 seconds
Minimum time between advertisement runs is 30 seconds
Received 13 messages, 0 notifications, 0 in queue
Sent 13 messages, 0 notifications, 0 in queue
Prefix advertised 1, suppressed 0, withdrawn 0
Connections established 1; dropped 0
Last reset never
1 accepted prefixes consume 32 bytes
0 history paths consume 0 bytes
External BGP neighbor may be up to 255 hops away
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-10

In the figure, the **show ip bgp neighbors** command has been executed on a router within a confederation. As a result, information about the intraconfederation EBGP session is displayed. The session is an external link (indicating an EBGP session) under common administration (indicating an intraconfederation EBGP session).


Monitoring Confederation Routes
Cisco.com
Route received from intra-confederation IBGP session
Next hop points to real EBGP peer in both cases
router# show ip bgp 14.0.0.0
BGP routing table entry for 14.0.0.0/8, version 5
Paths: (2 available, best #2, advertised over IBGP, EBGP)
(65001) 387
1.3.0.3 (metric 54357248) from 1.0.0.1 (11.0.0.1)
Origin IGP, metric 0, localpref 60, valid, confed-internal
(65001) 387
1.3.0.3 (metric 54357248) from 1.0.0.2 (10.1.1.1)
Origin IGP, metric 0, localpref 60, valid, confed-external, best
Intraconfederation part of AS path
External part of AS path
Route received from intra-confederation EBGP session
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-11

In this example, the **show ip bgp** command is executed on a router within the confederation to display information about the class A network 14.0.0.0.

The command response indicates that the router has received information about the network 14.0.0.0 on two different BGP sessions. One of the sessions is an intraconfederation EBGP session, and the other session is an IBGP session. Both updates have the same next-hop address, which was set by the true EBGP peer that originally sent the update into the confederation. The next hop is resolved by recursive routing, and, therefore, the forwarding decision will be the same regardless of which BGP entries are actually used. The second IP address is the address of the neighbor, which is followed by the router-ID of that neighbor and enclosed in parentheses.

The AS path is the same for both entries. It contains a parenthesized part, (65001). This is the part of the AS path that describes the intraconfederation AS path. The part of the AS path that follows is the external part, 387. This number reveals that the confederation has a true EBGP session with the official AS (AS 387), from which an update about network 14.0.0.0 was received. The update was forwarded to the router using IBGP within member-AS 65001. The router in the local member-AS, on which this command was executed, has two different intraconfederation EBGP sessions with member-AS 65001. So, the update about network 14.0.0.0 has entered the local AS via two different paths.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) How will the **show ip bgp** command display the intraconfederation segment of the AS path?

A) As a regular entry in the AS-path attribute.

B) As a separate AS-path list independent of the AS-path attribute.

C) As an entry in the AS path designated by parentheses.

D) The intraconfederation AS is not displayed because it is not a part of the EBGP AS path.

Q2) How does the **show ip bgp** command indicate that an EBGP session is with an intraconfederation peer instead of an actual EBGP peer?

A) by indicating that the neighbor is under common administration

B) by specifying that the session is over a confederation internal link

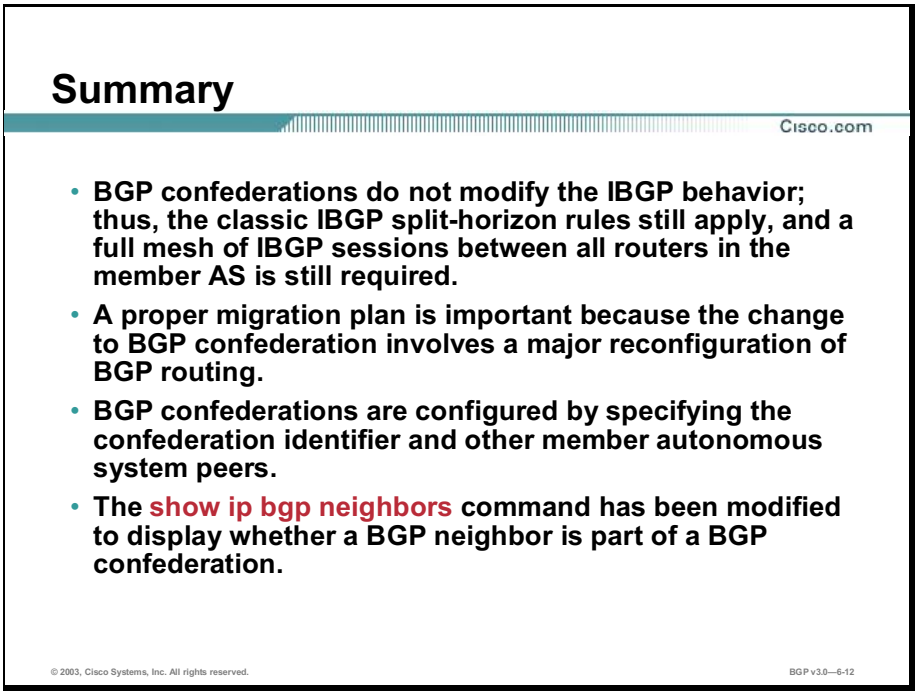
C) by displaying the next-hop address in parentheses

D) by indicating the BGP established state of the session to the external neighbor

Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• BGP confederations do not modify the IBGP behavior; thus, the classic IBGP split-horizon rules still apply, and a full mesh of IBGP sessions between all routers in the member AS is still required.
• A proper migration plan is important because the change to BGP confederation involves a major reconfiguration of BGP routing.
• BGP confederations are configured by specifying the confederation identifier and other member autonomous system peers.
• The show ip bgp neighbors command has been modified to display whether a BGP neighbor is part of a BGP confederation.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—6-12

Next Steps

After completing this lesson, go to:

- Optimizing BGP Scalability module

References

For additional information, refer to these resources:

- For more information on BGP confederations, refer to “Using the Border Gateway Protocol for Interdomain Routing” at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

- For more information on BGP confederations, refer to “Autonomous System Confederations for BGP” at the following URL: http://www.ietf.org/rfc/rfc1965.txt?number=1965

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: BGP Confederations

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10-BASE-T patch cables

Exercise Objective

In this exercise, you will enable confederations to modify IBGP AS-path processing in a transit AS.

After completing this exercise, you will be able to:

- Plan the migration of an existing IBGP backbone toward BGP confederations
- Configure BGP confederations on a BGP backbone
- Monitor a BGP backbone containing BGP confederations

Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
bgp confederation identifier *x*	Specify the real AS number.
bgp confederation peers *as1* [*as2* [*as3*] …]	List the neighboring intraconfederation autonomous systems.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.
neighbor *ip-address* **remote-as** *as-number*	Configure BGP neighbor.
neighbor *ip-address* **update-source** *interface*	Configure the source interface of a BGP session.
neighbor *ip-address* **ebgp-multihop**	Configure an EBGP session between neighbors that are not directly connected.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- The AS numbers that will be used in the confederation, as well as the BGP sessions that must be established, are shown in Figure 1.

- To successfully complete this exercise, you must remove all BGP processes and start new processes with private AS numbers according to the following rules:

Router	Member-AS Number
WGxR1	AS 65001
WGxR2, WGxR3	AS 65002
WGxR4	AS 65003

- Routers in member autonomous systems should originate their own networks. The sessions should be established on the routers between Loopback0 interfaces as shown in Figure 1.

- AS 99 should receive only prefixes originating in your AS x and AS 37. You should accept only prefixes originating in AS 99 from router “Client.”

 Copyright © 2003, Cisco Systems, Inc.

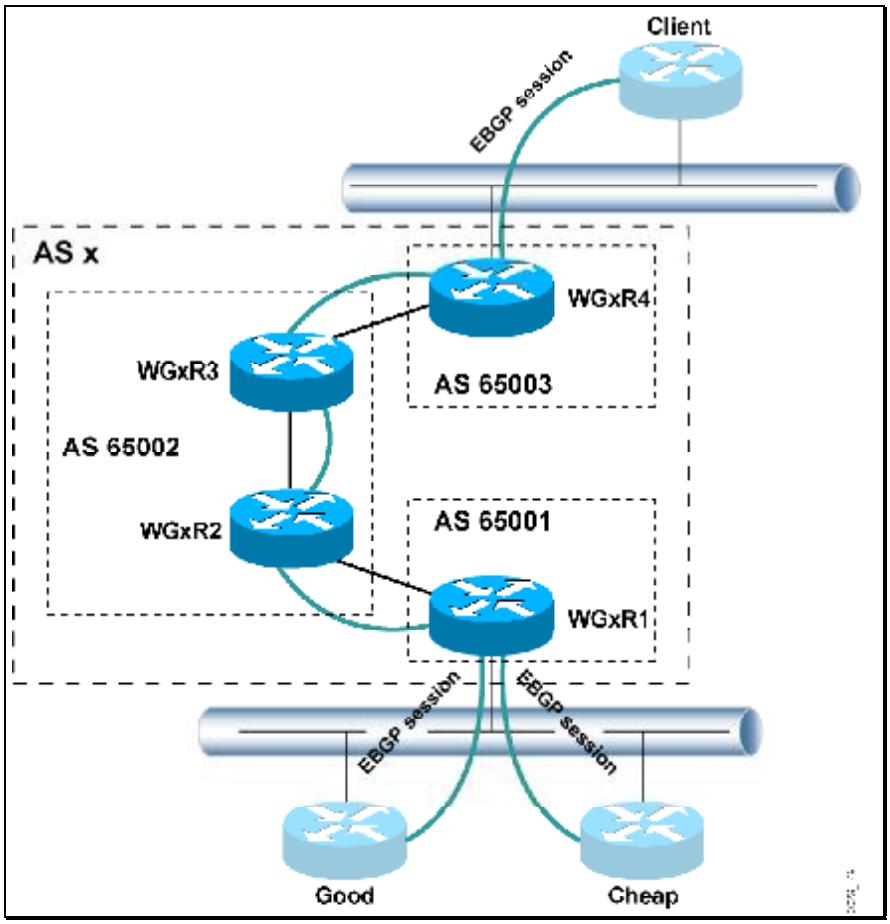


Figure 1: Implementing a BGP confederation design

Exercise Procedure

Complete these steps:

Perform the following steps on all routers in your workgroup:

Step 1 Remove the BGP process and start a new BGP process with member-AS number.

Step 2 Announce networks originating in this subautonomous system. Also announce network 192.168.x.0/24.

Step 3 Set the confederation identifier by using your real AS number. List all intraconfederation neighboring autonomous systems by using the **bgp confederation peers** command.

Step 4 Establish intraconfederation EBGP sessions and IBGP sessions within the member-AS. Use loopback interfaces for all BGP sessions.

Perform the following steps only on the specific router indicated in each step:

Step 5 On WGxR1, configure routers “Good” and “Cheap” as your BGP neighbors.

Step 6 On WGxR4, configure router “Client” as your BGP neighbor.

Step 7 On WGxR4, apply AS-path filters to incoming and outgoing updates toward router “Client.”

Copyright © 2003, Cisco Systems, Inc.

Note	The filter that permits networks originating in your AS needs to be changed to permit all prefixes originating in all intraconfederation autonomous systems.

Exercise Verification

You have completed this exercise when you attain these results:

- Check to see if router "Client" is receiving only your networks and those originating in AS 37.

```
Client#sh ip bgp regexp ^1_
BGP table version is 319, local router ID is 197.99.111.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 128.37.0.0       192.168.21.1                           0 1 20 42
37 i
*> 192.37.11.0      192.168.21.1                           0 1 20 42
37 i
*> 192.168.1.0      192.168.21.1             0             0 1 i
*> 197.1.1.0        192.168.21.1                           0 1 i
*> 197.1.2.0        192.168.21.1                           0 1 i
*> 197.1.3.0        192.168.21.1                           0 1 i
*> 197.1.4.0        192.168.21.1                           0 1 i
*> 197.1.5.0        192.168.21.1                           0 1 i
*> 197.1.6.0        192.168.21.1             0             0 1 i
*> 197.1.7.0        192.168.21.1             0             0 1 i
*> 197.1.8.0        192.168.21.1                           0 1 i
```

- Use the **show ip route ospf** command to make sure that your IGP carries only your internal networks.

Answer these questions:

Q1) What additional command did you have to use to establish intraconfederation EBGP sessions between loopback interfaces?

Q2) Why was it necessary to change the AS-path filters?

 Copyright © 2003, Cisco Systems, Inc.

7

Optimizing BGP Scalability

Overview

The Border Gateway Protocol (BGP) is designed for reliability and scalability. As such, it has become the de facto standard protocol used to carry more than 110,000 prefixes in the Internet today. Likewise, BGP has a tremendous amount of flexibility with regard to administrative policy controls, route selection, performance tuning, and scalability features. This module introduces advanced BGP configuration tools designed to improve BGP scalability and performance. Tools discussed in this module include convergence time reduction features; limiting the number of prefixes; peer groups; and route dampening.

Upon completing this module, you will be able to:

- Configure Cisco IOS® performance improvements to reduce BGP convergence time, given a typical BGP network
- Successfully configure BGP to limit the number of prefixes received from a neighbor, given a typical BGP network
- Use BGP peer groups to share common configuration parameters between multiple BGP peers, given a typical BGP network
- Use route dampening to minimize the impact of unstable routes, given a properly configured BGP network

Outline

The module contains these lessons:

- Improving BGP Convergence
- Limiting the Number of Prefixes Received from a BGP Neighbor
- BGP Peer Groups
- BGP Route Dampening

 Copyright © 2003, Cisco Systems, Inc.

Improving BGP Convergence

Overview

Border Gateway Protocol (BGP) is the routing protocol used to communicate between networks connected to the Internet. As the Internet routing table grows, service providers and large enterprise customers are experiencing a dramatic increase in the amount of time that BGP takes to converge. Networks that once converged in 10 or 15 minutes may now take up to one hour in some cases, and even longer in extreme situations.

Several BGP enhancements have been made in Cisco IOS® to improve convergence and basic scaling properties. In addition, there are several steps that customers can take to further reduce BGP convergence times. This lesson introduces different Cisco IOS performance improvements designed to reduce BGP convergence time. Included in this lesson are discussions of convergence, BGP routing processes, and the effects of BGP routing processes on router CPU resources. The lesson also discuses the commands required to configure and monitor BGP for different Cisco IOS performance improvements, including: path maximum transmission unit (MTU) discovery, increasing the input hold queue, BGP scan time, and the BGP advertisement interval.

Importance

As the number of routes in the Internet increases, demands on router CPU and memory resources will increase. BGP processing affects both router resources and network convergence time. It is important that network convergence be as fast as possible to ensure accurate routing information between domains. It is also important that router resources be optimized whenever possible. Cisco IOS performance improvements for BGP are designed to aid network administrators in achieving these goals.

Objectives

Upon completing this lesson, you will be able to:

- Describe convergence in BGP networks
- Describe the function of BGP router processes
- Describe the effects of BGP processes on router CPU resources
- Identify Cisco IOS performance improvements to reduce BGP convergence time
- Identify the Cisco IOS commands required to configure and monitor path MTU discovery
- Identify the Cisco IOS commands required to configure and monitor the input queue depth on a router interface
- Identify the Cisco IOS commands required to configure and monitor BGP scan time
- Identify the Cisco IOS commands required to configure and monitor the BGP advertisement interval

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

 Copyright © 2003, Cisco Systems, Inc.

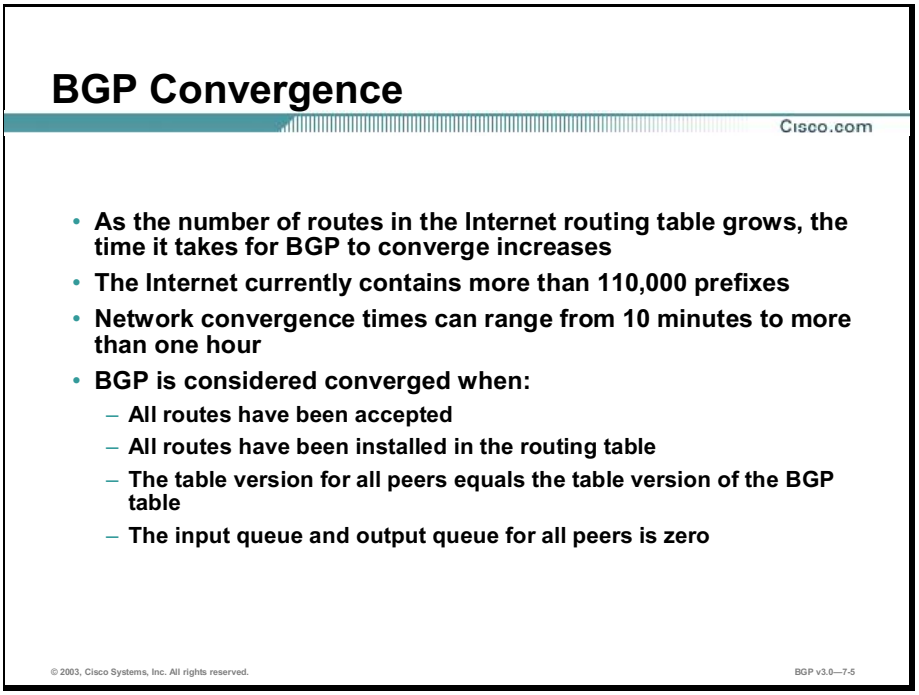
Outline

This lesson includes these topics:

- Overview
- BGP Convergence
- BGP Processes
- CPU Effects of BGP Processes
- Improving BGP Convergence
- Path MTU Discovery
- Increasing Input Queue Depth
- BGP Scan Time
- BGP Advertisement Interval
- Summary
- Assessment (Quiz): Improving BGP Convergence

BGP Convergence

This topic describes the concept of convergence in BGP networks.


BGP Convergence
Cisco.com
• As the number of routes in the Internet routing table grows, the time it takes for BGP to converge increases
• The Internet currently contains more than 110,000 prefixes
• Network convergence times can range from 10 minutes to more than one hour
• BGP is considered converged when:
– All routes have been accepted
– All routes have been installed in the routing table
– The table version for all peers equals the table version of the BGP table
– The input queue and output queue for all peers is zero
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-5

As the number of routes in the Internet routing table grows, so too does the time it takes for BGP to converge. In general, convergence is defined as the process of bringing all route tables to a state of consistency. The BGP routing protocol is considered "converged" when the following conditions are true:

- All routes have been accepted.
- All routes have been installed in the routing table.
- The table version for all peers equals the table version of the BGP table.
- The input queue and output queue for all peers is 0.

Convergence time is an important consideration in a network, as non-converged networks can cause routing loops, packet delays, and even packet loss due to "black holes."

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) Why is the speed at which a routing protocol converges important?

A) A nonconverged routing protocol cannot forward packets.

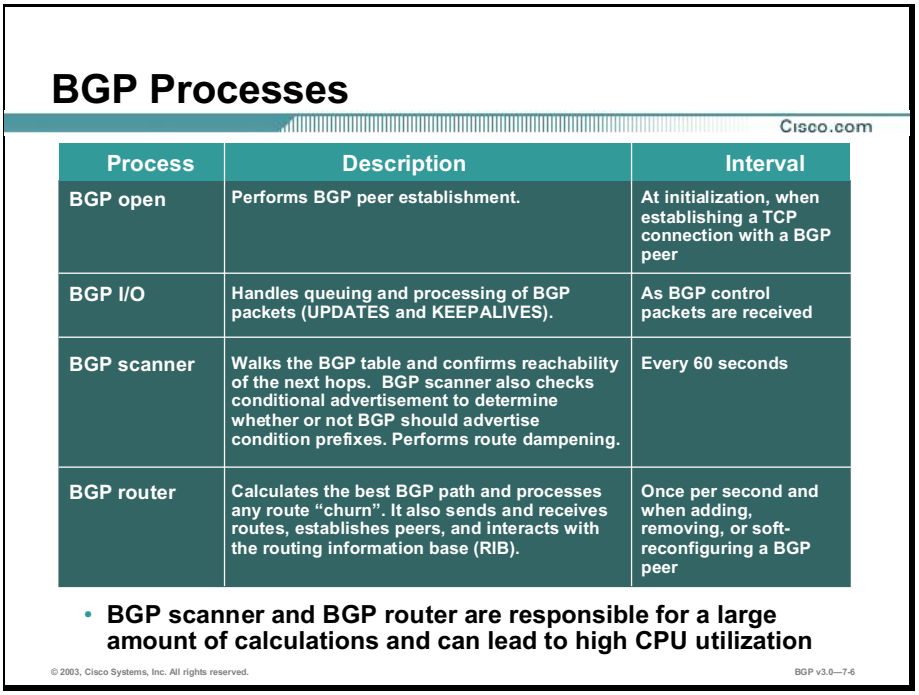
B) Routing metrics cannot be properly computed unless the network is converged.

C) Routing algorithms that converge slowly can cause routing loops or network outages.

D) Networks that are not converged have reduced security.

BGP Processes

This topic describes the different BGP router processes and their function.


BGP Processes
Cisco.com
Process
Description
Interval
BGP open
Performs BGP peer establishment.
At initialization, when establishing a TCP connection with a BGP peer
BGP I/O
Handles queuing and processing of BGP packets (UPDATES and KEEPALIVES).
As BGP control packets are received
BGP scanner
Walks the BGP table and confirms reachability of the next hops. BGP scanner also checks conditional advertisement to determine whether or not BGP should advertise condition prefixes. Performs route dampening.
Every 60 seconds
BGP router
Calculates the best BGP path and processes any route "churn". It also sends and receives routes, establishes peers, and interacts with the routing information base (RIB).
Once per second and when adding, removing, or soft-reconfiguring a BGP peer
• BGP scanner and BGP router are responsible for a large amount of calculations and can lead to high CPU utilization
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-6

In general, a Cisco IOS process consists of the individual threads and associated data that perform router tasks, such as system maintenance, packet switching, and implementing routing protocols.

Reference A thread is an information placeholder that allows a single process to be halted (interrupted) on the router so that the CPU can service another process. The information contained within the thread allows the interrupted process to restart exactly where it left off once the CPU is ready to service that process thread once again.

Several Cisco IOS processes executed on the router enable BGP to run. You can use the **show process cpu | include BGP** command to see the amount of CPU resources consumed (utilization) due to running BGP processes.

The figure lists the function of each of the BGP router processes and how often each process is executed on the router. It shows that each process runs at different times depending on the tasks that are handled by the specific process. Because BGP scanner and BGP router are responsible for a large amount of calculations, you may notice high CPU utilization during the running of either one of these processes.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What two BGP processes can consume large amounts of CPU resources? (Choose two.)

A) BGP open

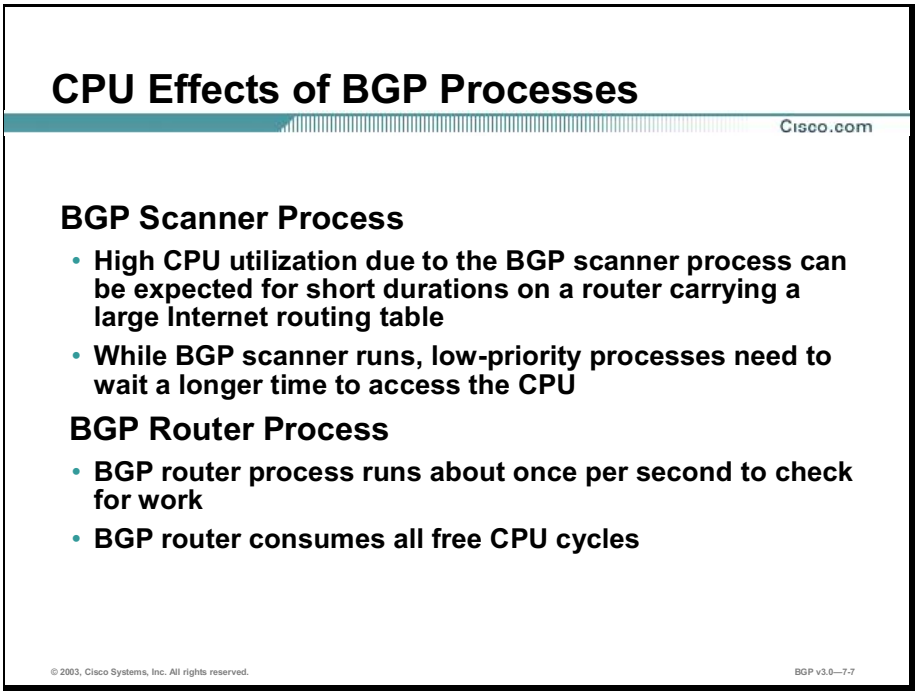
B) BGP scanner

C) BGP I/O

D) BGP router

CPU Effects of BGP Processes

This topic describes how running BGP router processes affect router CPU resources.


CPU Effects of BGP Processes
Cisco.com
BGP Scanner Process
• High CPU utilization due to the BGP scanner process can be expected for short durations on a router carrying a large Internet routing table
• While BGP scanner runs, low-priority processes need to wait a longer time to access the CPU
BGP Router Process
• BGP router process runs about once per second to check for work
• BGP router consumes all free CPU cycles
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-7

On routers carrying a large Internet routing table, high CPU utilization due to the BGP scanner process can be expected for short periods of time. Once per minute, BGP scanner walks (scans) the BGP routing information base (RIB) table and performs important maintenance tasks. These tasks include checking the next hop referenced in the BGP table of the router and verifying that the next-hop devices can be reached. Thus, a large BGP table takes an equally large amount of time to be walked and validated.

BGP scanner walks the BGP table to update any data structures and walks the routing table for route redistribution purposes. In this context, the routing table is also known as the RIB, which the router outputs when the **show ip route** command is executed. Both tables are stored separately in the router memory and can be very large, thus consuming CPU and memory resources.

Because BGP scanner runs through the entire BGP table, the duration of the high CPU utilization condition caused by the BGP scanner process varies with the number of neighbors and the number of routes learned per neighbor.

While BGP scanner runs, low-priority processes need to wait a longer time to access the CPU. One low-priority process controls Internet Control Message Protocol (ICMP) packets such as pings. Packets destined to or originated from the router may experience higher than expected latency because the ICMP process must wait behind BGP scanner. The BGP scanner process runs for some time, is suspended, then ICMP runs, ICMP is suspended, BGP scanner runs, and so on. In contrast, pings sent through a router should be switched via Cisco Express Forwarding (CEF) and should not experience any additional latency. When troubleshooting periodic spikes

 Copyright © 2003, Cisco Systems, Inc.

in latency, compare forwarding times for packets forwarded through a router versus packets processed directly by the CPU on the router.

The BGP router process runs about once per second to check for work. BGP convergence defines the duration between the time when the first BGP peer is established and the point at which BGP is converged. To ensure the shortest possible convergence times, BGP router consumes all free CPU cycles. However, after it starts, it relinquishes (or suspends) the CPU intermittently.

Example

Convergence time is a direct measurement of how long BGP router process runs on the CPU, not the total time that the process is actually running. This example investigates the high CPU utilization condition during BGP convergence as BGP exchanges prefixes with two External Border Gateway Protocol (EBGP) peers.

Step 1 Capture a baseline for normal CPU utilization before starting the test.

```
router# show process cpu

CPU utilization for five seconds: 0%/0%; one minute: 4%; five
minutes: 5%
```

Step 2 Once the test starts, the CPU reaches 100 percent utilization. The **show process cpu** command shows that the high CPU condition is caused by BGP router, denoted by 139 (the Cisco IOS process ID for BGP router) in the following output.

```
router# show process cpu

CPU utilization for five seconds: 100%/0%; one minute: 99%; five
minutes:

81%[output omitted] 139     6795740   1020252       6660 88.34%
91.63% 74.01%   0 BGP Router
```

Step 3 Monitor the router by capturing multiple outputs of the **show ip bgp summary** and **show process cpu** commands during the event. The **show ip bgp summary** command captures the state of the BGP neighbors.

```
router# show ip bgp summary

Neighbor    V    AS  MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd

10.1.1.1    4  64512  309453  157389    19981    0  253 22:06:44
111633

172.16.1.1  4  65101  188934    1047    40081   41    0 00:07:51
58430
```

Step 4 When the router completes prefix exchange with its BGP peers, the CPU utilization rates should return to normal levels. The computed 1-minute and 5-minute averages will settle back down as well and may show higher than normal levels for a longer period than the 5-second rate.

```
router# show process cpu

CPU utilization for five seconds: 3%/0%; one minute: 82%; five
minutes: 91%
```

Step 5 Using the output from the above **show** commands will allow you to compute the BGP convergence time. In particular, the "Up/Down" column of the **show ip bgp summary** command is compared to the start and stop times of the high CPU utilization condition. Typically, BGP convergence can take several minutes when routers exchange a large Internet routing table.

Practice

Q1) Why are BGP filtering tools such as AS-path filters and prefix-lists important in managing CPU resources in BGP-enabled routers?

A) Because filtering tools make the router inherently more secure.

B) These tools help reduce the size of the BGP table that is walked by BGP processes.

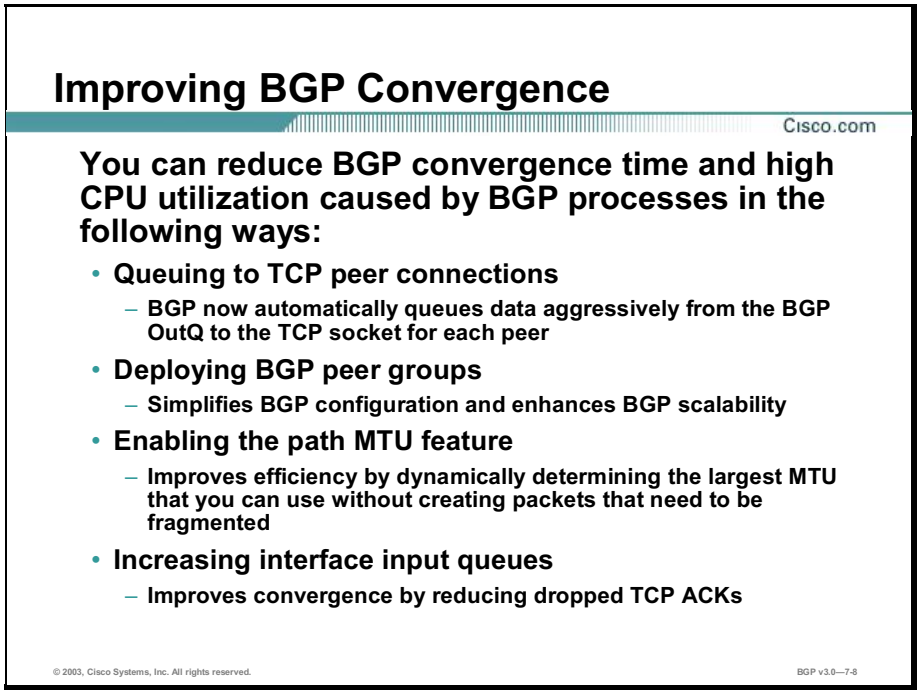
C) These tools switch BGP processing into CEF switching mode.

D) These tools reduce the number of BGP attributes attached to BGP routes.

 Copyright © 2003, Cisco Systems, Inc.

Improving BGP Convergence

This topic identifies Cisco IOS performance improvements that reduce BGP convergence times.


Improving BGP Convergence
Cisco.com
You can reduce BGP convergence time and high CPU utilization caused by BGP processes in the following ways:
• Queuing to TCP peer connections
– BGP now automatically queues data aggressively from the BGP OutQ to the TCP socket for each peer
• Deploying BGP peer groups
– Simplifies BGP configuration and enhances BGP scalability
• Enabling the path MTU feature
– Improves efficiency by dynamically determining the largest MTU that you can use without creating packets that need to be fragmented
• Increasing interface input queues
– Improves convergence by reducing dropped TCP ACKs
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-8

BGP convergence can often be an issue in networks, requiring quick propagation of routing information. Cisco IOS software provides the following performance-improvement features designed to reduce BGP convergence time and the high CPU utilization caused by a running BGP process.

- **Queuing to TCP peer connections**: Instead of queuing data once per second, BGP now queues data aggressively from the BGP output queue to the TCP socket for each peer until the output queues have drained completely. Because BGP now sends at a faster rate, it converges more quickly.

- **BGP peer groups**: The major benefit of specifying a BGP peer group is that it reduces the amount of system resources (CPU and memory) used in BGP update generation. Peer groups also simplify BGP configuration because many repetitive configuration elements (such as filters) are applied by the router only once (to the peer group) instead of applying them to each neighbor.

  Because peer groups allow the routing table to be checked only once, and allow updates to be replicated to all other in-sync peer group members (depending on the number of peer group members, the number of prefixes in the table, and the number of prefixes advertised), they can significantly reduce router resource requirements.

- **Path MTU feature**: All TCP sessions are bounded by a limit on the number of bytes that a single packet can transport. This limit, known as the Maximum Segment Size (MSS), is

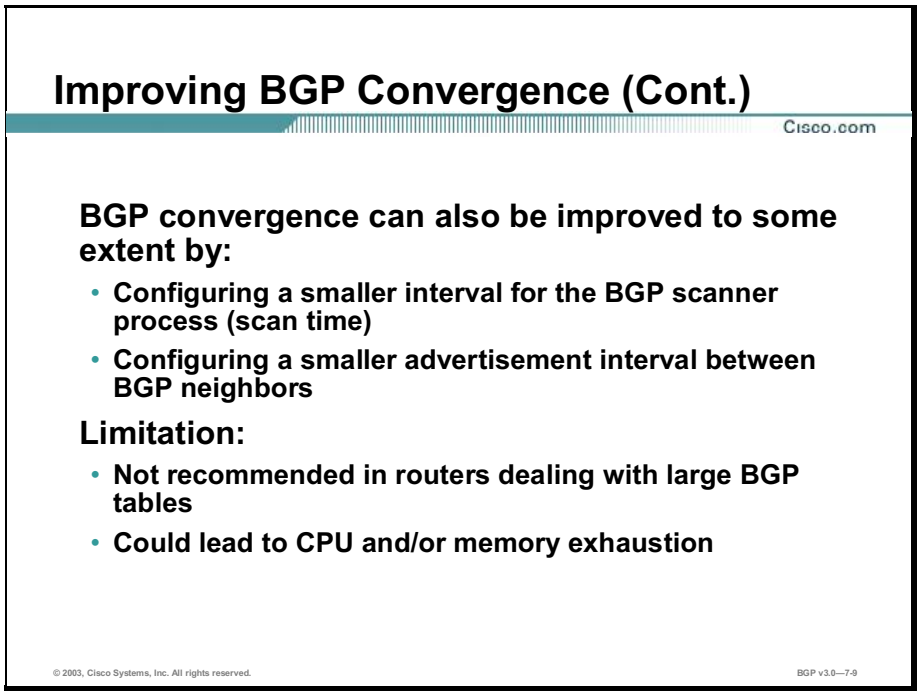
536 bytes by default. In other words, TCP breaks up packets in a transmit queue into 536-byte chunks before passing packets down to the IP layer. The advantage of a 536-byte MSS is that packets are not likely to be fragmented at an IP device along the path to the destination, because most links use an MTU of at least 1500 bytes. The disadvantage is that smaller packets increase the amount of bandwidth used to transport overhead.

Because BGP builds a TCP connection to all peers, a 536-byte MSS affects BGP convergence times. The solution is to enable the path maximum transmission unit (PMTU) feature by means of the **ip tcp path-mtu-discovery** command. You can use this feature to dynamically determine how large the MSS value can be without creating packets that need to be fragmented. PMTU allows TCP to determine the smallest MTU size among all links in a TCP session. TCP then uses this MTU value, minus room for the IP and TCP headers, as the MSS for the session.

- **Increase interface input queues:** If BGP is advertising thousands of routes to many neighbors, TCP must transmit thousands of packets. BGP peers receive these packets and send TCP acknowledgments (ACKs) to the advertising BGP speaker, causing the BGP speaker to receive a flood of TCP ACKs in a short period of time. If the ACKs arrive at a rate that is too high for the router CPU, packets back up in inbound interface queues.

  By default, router interfaces use an input queue size of 75 packets. In addition, special control packets such as BGP UPDATES use a special queue with Selective Packet Discard (SPD). This special queue holds 100 packets. During BGP convergence, TCP ACKs can quickly fill the 175 spots of input buffering, causing newly arriving packets to be dropped. On routers with 15 or more BGP peers that also exchange the full Internet routing table, over 10,000 drops per interface per minute may be seen. Increasing the interface input queue depth using the **hold-queue in** command helps reduce the number of dropped TCP ACKs, thereby reducing the amount of work that BGP must do to converge.

 Copyright © 2003, Cisco Systems, Inc.


Improving BGP Convergence (Cont.)
Cisco.com
BGP convergence can also be improved to some extent by:
• Configuring a smaller interval for the BGP scanner process (scan time)
• Configuring a smaller advertisement interval between BGP neighbors
Limitation:
• Not recommended in routers dealing with large BGP tables
• Could lead to CPU and/or memory exhaustion
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-9

Network administrators also need to improve BGP convergence in certain scenarios; for example, in networks using the conditional advertisement feature. There are two additional BGP parameters that they can use to influence BGP convergence speed:

- **Scan time:** Controlling the BGP scanner process, responsible for verifying information in the BGP table
- **Advertisement interval:** Controlling the rate at which successive advertisements are sent to a BGP neighbor

Network administrators must take care when configuring these two parameters. Setting the values too low for a specific network environment could lead to a significant consumption of router resources. The larger the BGP tables and the more unstable the BGP network, the greater the danger of exhausting the resources of a router.

Practice

Q1) What BGP convergence-enhancing feature reduces convergence by first reducing transport overhead and processing delays at network transit points?

A) aggressive TCP packet queuing

B) PMTU discovery

C) scan time

D) peer groups

Q2) What two BGP convergence-enhancing tools are recommended for all BGP implementations? (Choose two.)

A) PMTU discovery

B) scan time

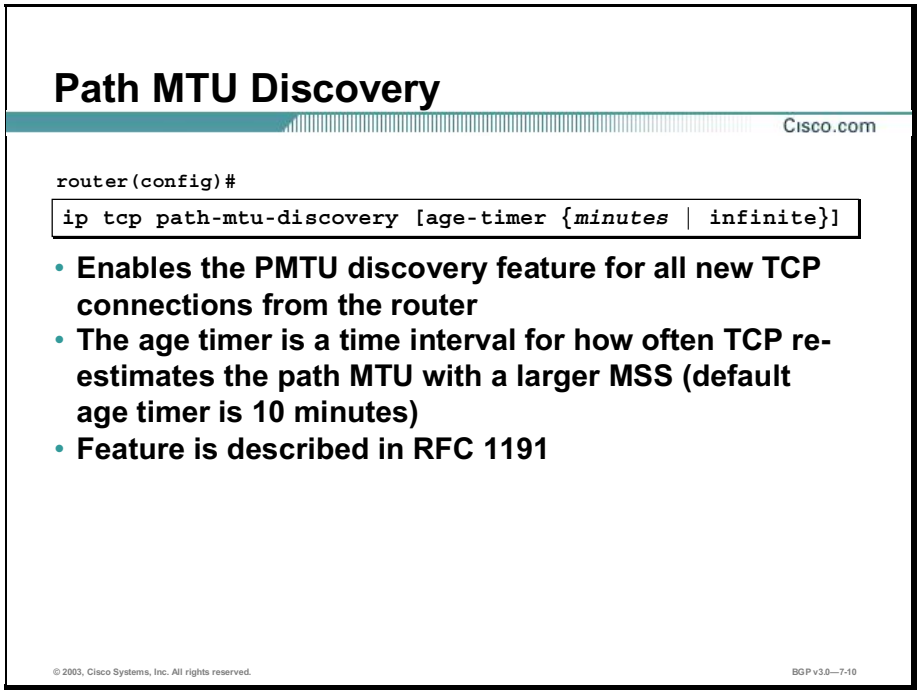
C) advertisement interval

D) peer groups

 Copyright © 2003, Cisco Systems, Inc.

Path MTU Discovery

This topic identifies the Cisco IOS commands required to configure and monitor PMTU discovery.


Path MTU Discovery
Cisco.com
router(config)#
ip tcp path-mtu-discovery [age-timer {minutes | infinite}]
• Enables the PMTU discovery feature for all new TCP connections from the router
• The age timer is a time interval for how often TCP re-estimates the path MTU with a larger MSS (default age timer is 10 minutes)
• Feature is described in RFC 1191
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-10

PMTU discovery is a method for maximizing the use of available bandwidth in the network between the endpoints of a TCP connection. It is described in RFC 1191. Existing connections are not affected when this feature is turned on or off.

Customers using TCP connections to move bulk data between systems on distinct subnets would benefit most by enabling this feature.

The age timer is a time interval for how often TCP re-estimates the PMTU with a larger Maximum Segment Size (MSS). The default value of the age timer is 10 minutes, but it can be manually configured up to 30 minutes or disabled (set to infinite). If the MSS used for the connection is smaller than what the peer connection can handle, the router will attempt to use a larger MSS each time that the age timer expires. The discovery process is stopped when either the send MSS is as large as the peer negotiated, or the user has disabled the timer on the router. You can turn off the age timer by setting it to “infinite.”

ip tcp path-mtu-discovery

To enable the PMTU discovery feature for all new TCP connections from the router, use the **ip tcp path-mtu-discovery** global configuration command.

ip tcp path-mtu-discovery [**age-timer** {*minutes* | **infinite**}]

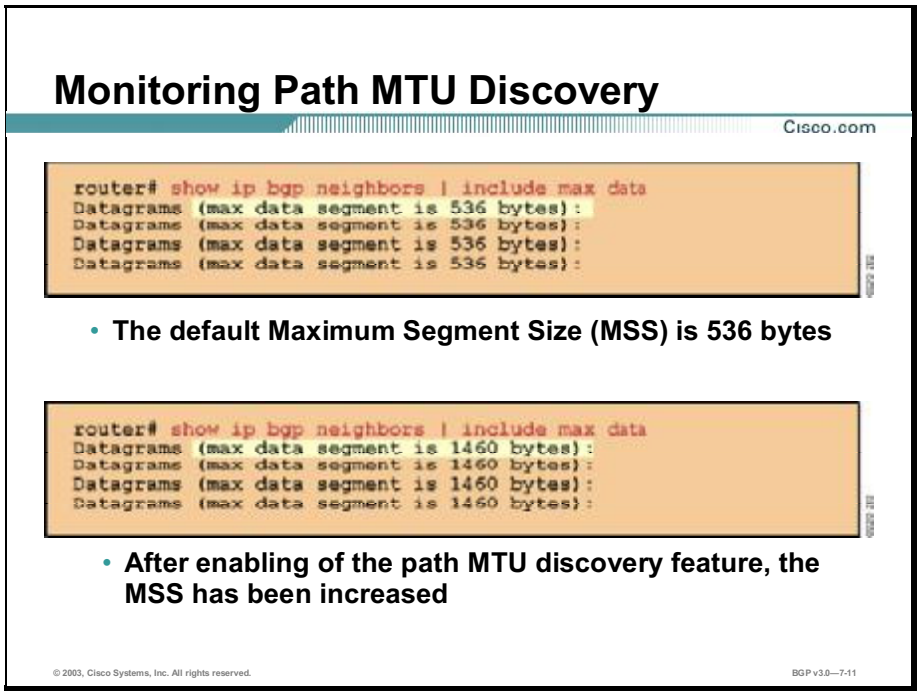
To disable the function, use the **no** form of this command.

no ip tcp path-mtu-discovery [**age-timer** {*minutes* | **infinite**}]

Syntax Description

age-timer *minutes*	(Optional) Time interval (in minutes) after which TCP re-estimates the PMTU with a larger MSS. The maximum interval is 30 minutes; the default is 10 minutes.
age-timer infinite	(Optional) Turns off the age timer.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Path MTU Discovery
Cisco.com
router# show ip bgp neighbors | include max data
Datagrams (max data segment is 536 bytes):
Datagrams (max data segment is 536 bytes):
Datagrams (max data segment is 536 bytes):
Datagrams (max data segment is 536 bytes):
• The default Maximum Segment Size (MSS) is 536 bytes
router# show ip bgp neighbors | include max data
Datagrams (max data segment is 1460 bytes):
Datagrams (max data segment is 1460 bytes):
Datagrams (max data segment is 1460 bytes):
Datagrams (max data segment is 1460 bytes):
• After enabling of the path MTU discovery feature, the MSS has been increased
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-11

By default, the MSS is 536 bytes. As shown in the figure, the **show ip bgp neighbors | include max data** command can be used to verify the size of the MSS before the PMTU discovery feature is enabled on the router.

After using the **ip tcp path-mtu-discovery** command to enable PMTU discovery, the router dynamically determines how large the MSS can be without creating IP packets that require fragmentation. At the bottom of figure, the output shows that the PMTU feature has been enabled and the **show ip bgp neighbors | include max data** command has been used to determine that the PMTU discovery feature has set the MSS to 1460 bytes.

Practice

Q1) In what two ways does the PMTU discovery feature aid in improving BGP convergence? (Choose two.)

A) It allows the router to determine the MSS of the network path.

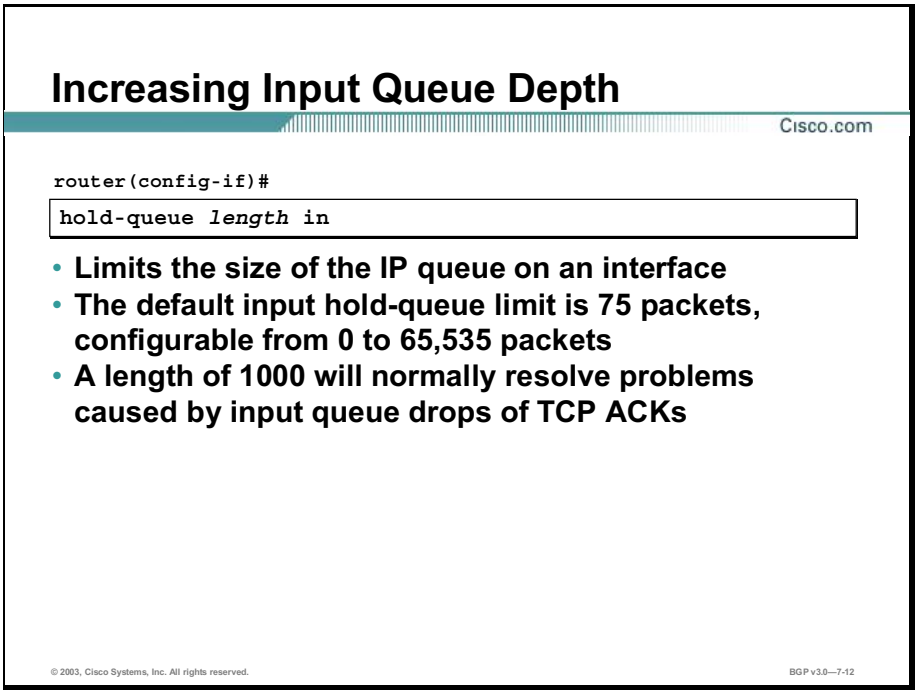
B) It allows the router to send TCP packets only across network paths with the highest MSS.

C) It reduces the amount of overhead in transporting data between TCP sessions.

D) It removes the increased CPU and memory overhead required to fragment an IP datagram.

Increasing Input Queue Depth

This topic identifies the Cisco IOS commands required to configure and monitor the input queue depth on a router interface.


Increasing Input Queue Depth
Cisco.com
router(config-if)#
hold-queue length in
• Limits the size of the IP queue on an interface
• The default input hold-queue limit is 75 packets, configurable from 0 to 65,535 packets
• A length of 1000 will normally resolve problems caused by input queue drops of TCP ACKs
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-12

Each interface owns an input queue onto which incoming packets are placed to await processing by the router. Frequently, the rate of incoming packets placed on the input queue exceeds the rate at which the router can process the packets. Each input queue has a size that indicates the maximum number of packets that may be placed in the queue. After the input queue becomes full, the interface drops any new incoming packets.

hold-queue

To specify the size of the IP output queue on an interface, use the **hold-queue** command in interface configuration mode.

hold-queue *length* {**in** | **out**}

To restore the default values for an interface, use the **no** form of this command with the appropriate keyword.

no hold-queue *length* {**in** | **out**}

Syntax Description

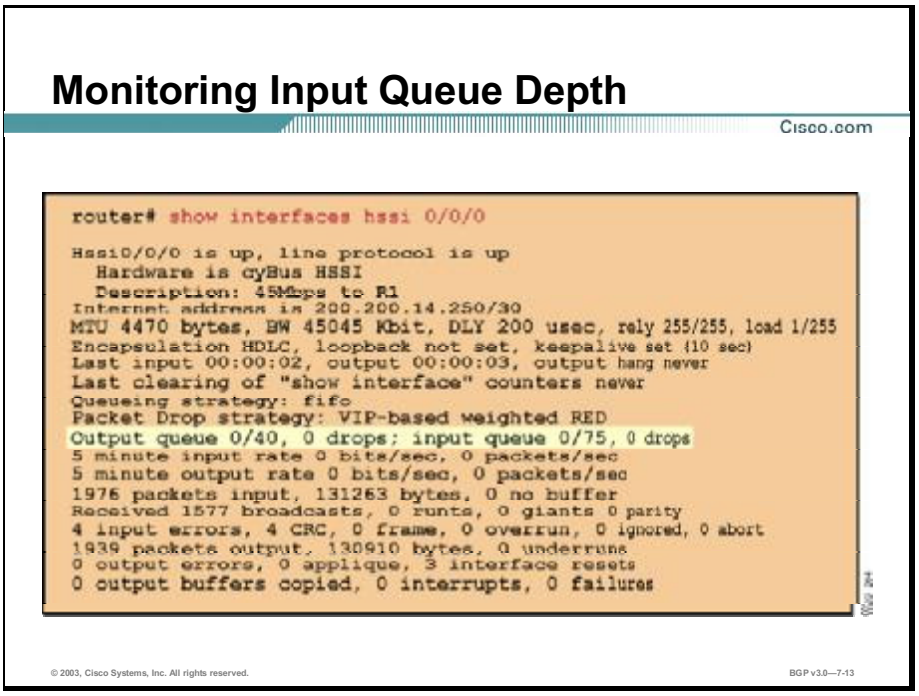
length	Integer that specifies the maximum number of packets in the queue. The range of allowed values is 0 to 65535.

 Copyright © 2003, Cisco Systems, Inc.

in	Specifies the input queue. The default is 75 packets. For asynchronous interfaces, the default is 10 packets. These limits prevent a malfunctioning interface from consuming an excessive amount of memory.
out	Specifies the output queue. The default is 40 packets. For asynchronous interfaces, the default is 10 packets. These limits prevent a malfunctioning interface from consuming an excessive amount of memory.

Caution Increasing the hold queue can have detrimental effects on network routing and response times. For protocols that use SEQ/ACK packets to determine round-trip times, do not increase the output queue. Dropping packets instead informs hosts to slow down transmissions to match available bandwidth. This approach is generally better than having duplicate copies of the same packet within the network (which can happen with large hold queues).

Note The Cisco 12000 Series now uses a default SPD headroom value of 1000. It retains the default input queue size of 75. Use the **show spd** command to view these special input queues.


Monitoring Input Queue Depth
Cisco.com
router# show interfaces hssi 0/0/0
Hssi0/0/0 is up, line protocol is up
Hardware is cyBus HSSI
Description: 45Mbps to R1
Internet address is 200.200.14.250/30
MTU 4470 bytes, BW 45045 Kbit, DLY 200 usec, rely 255/255, load 1/255
Encapsulation HDLC, loopback not set, keepalive set (10 sec)
Last input 00:00:02, output 00:00:03, output hang never
Last clearing of "show interface" counters never
Queueing strategy: fifo
Packet Drop strategy: VIP-based weighted RED
Output queue 0/40, 0 drops; input queue 0/75, 0 drops
5 minute input rate 0 bits/sec, 0 packets/sec
5 minute output rate 0 bits/sec, 0 packets/sec
1976 packets input, 131263 bytes, 0 no buffer
Received 1577 broadcasts, 0 runts, 0 giants 0 parity
4 input errors, 4 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort
1939 packets output, 130910 bytes, 0 underruns
0 output errors, 0 applique, 3 interface resets
0 output buffers copied, 0 interrupts, 0 failures
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-13

In the figure, the **show interface** {*interface-identifier*} command displays the current input queue levels and the number of incoming packets dropped.

The input queue x/y counter displays the current number of packets in the input queue “x” and the current size of the input queue “y”. The drops counter indicates the number of incoming packets that have been dropped.

If the current number of packets in the input queue is consistently at or greater than 80 percent of the current size of the input queue, the size of the input queue may require tuning to accommodate the rate of incoming packets. Even if the current number of packets on the input queue never seems to approach the size of the input queue, bursts of packets may still be overflowing the queue. If the drops counter is increasing at a high rate, the size of the input queue may require tuning to accommodate the bursts.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) How does increasing the size of the input queue on an interface aid in improving BGP convergence?

A) It allows for larger packets to be sent between BGP neighbors.

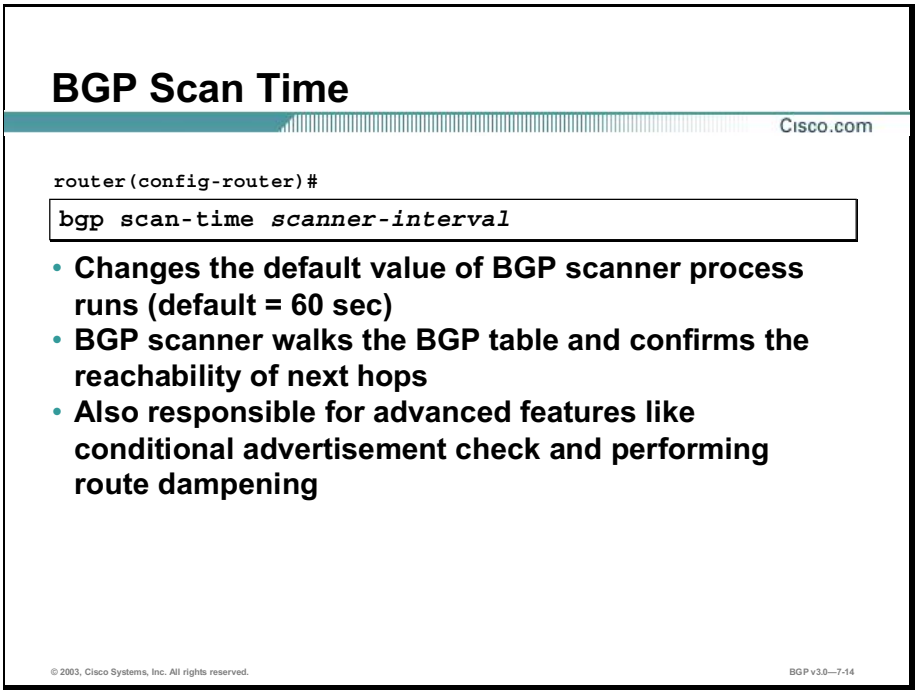
B) It increases the number of BGP updates that can be contained in a single TCP segment.

C) It reduces overhead in processing BGP update packets by reducing the number of retransmissions caused by dropped TCP ACK packets.

D) It increases the rate at which incoming packets can be sent to the output queue of the outgoing interface.

BGP Scan Time

This topic identifies the Cisco IOS commands required to configure and monitor BGP scan time.


BGP Scan Time
Cisco.com
router(config-router)#
bgp scan-time scanner-interval
• Changes the default value of BGP scanner process runs (default = 60 sec)
• BGP scanner walks the BGP table and confirms the reachability of next hops
• Also responsible for advanced features like conditional advertisement check and performing route dampening
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-14

Network administrators use the **bgp scan-time** command to configure the time interval for repetitions of the BGP scanner process.

Recall that the BGP scanner process walks (scans) the BGP table and confirms the reachability of next hops. A change of this status will trigger a new BGP route selection for the network. Changes are then propagated by the router to established BGP neighbors. Increasing the BGP scanner process frequency will make the router find a changed status more quickly, but it will also consume more CPU resources.

The BGP scanner process is also responsible for some advanced BGP features. It checks the conditional advertisement to determine whether or not BGP should advertise conditional prefixes or perform route dampening.

bgp scan-time

To configure a nondefault value of the scanning interval for BGP routing information, use the **bgp scan-time** command.

bgp scan-time [**import**] *scanner-interval*

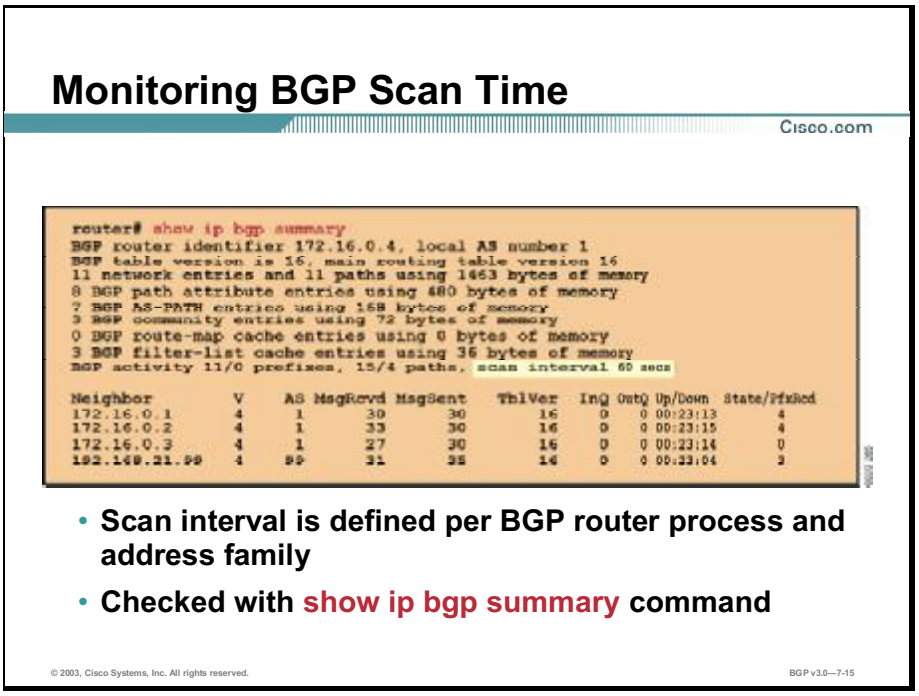
To disable the scanning interval of a router, use the **no** form of this command.

no bgp scan-time

 Copyright © 2003, Cisco Systems, Inc.

Syntax Description

import	(Optional) Configures import processing of VPNv4 unicast routing information from BGP routers into routing tables.
scanner-interval	Specifies the scanning interval of BGP routing information. Valid values used for selecting the desired scanning interval are from 5 to 60 seconds. By default, the scanning interval is 60 seconds.


Monitoring BGP Scan Time
Cisco.com
router# show ip bgp summary
BGP router identifier 172.16.0.4, local AS number 1
BGP table version is 16, main routing table version 16
11 network entries and 11 paths using 1463 bytes of memory
8 BGP path attribute entries using 480 bytes of memory
7 BGP AS-PATH entries using 168 bytes of memory
3 BGP community entries using 72 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
3 BGP filter-list cache entries using 36 bytes of memory
BGP activity 11/0 prefixes, 15/4 paths, scan interval 60 secs
Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd
172.16.0.1 4 1 30 30 16 0 0 00:23:13 4
172.16.0.2 4 1 33 30 16 0 0 00:23:15 4
172.16.0.3 4 1 27 30 16 0 0 00:23:14 0
192.168.21.99 4 99 31 35 16 0 0 00:23:04 3
• Scan interval is defined per BGP router process and address family
• Checked with show ip bgp summary command
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-15

You can check the configured BGP scan interval using the **show ip bgp summary** command.

The configured BGP scan interval will apply to the entire BGP routing protocol process.

Practice

Q1) What negative effects can occur when you are reducing the scan time in an attempt to improve BGP convergence?

A) Frequent runs of the BGP scanner process may cause BGP to converge before the IGP does, resulting in network black holes.

B) If the scanner process runs too frequently, the router CPU and memory resources may become overwhelmed.

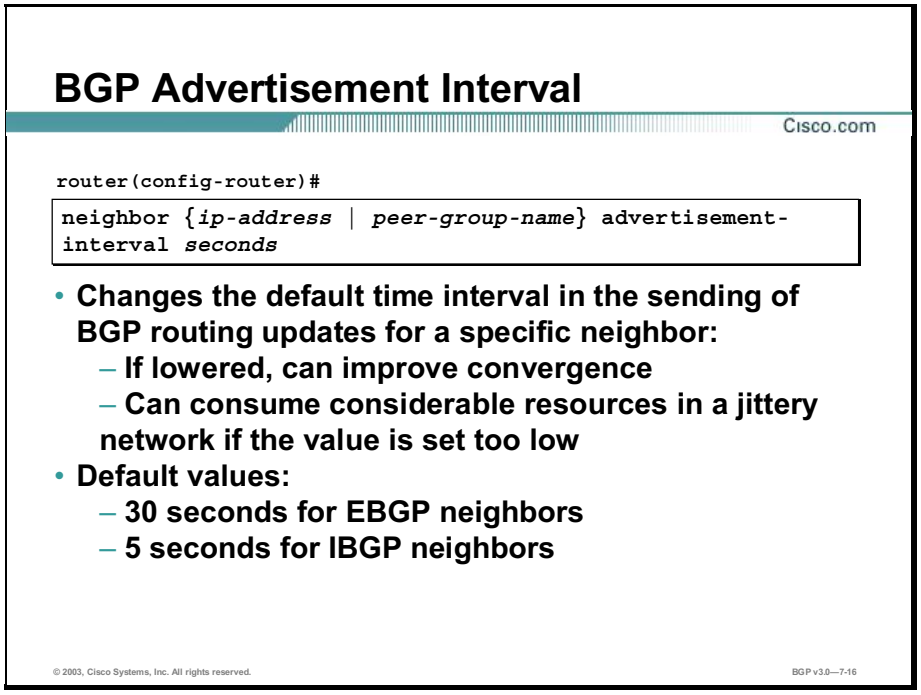
C) If the BGP scanner process runs too often, the BGP router process will not get enough CPU cycles to properly update the routing table.

D) If the BGP scanner process runs too often, the BGP open process will not get enough CPU cycles to properly maintain BGP neighbor relationships.

 Copyright © 2003, Cisco Systems, Inc.

BGP Advertisement Interval

This topic identifies the Cisco IOS commands required to configure and monitor the BGP advertisement interval.


BGP Advertisement Interval
Cisco.com
router(config-router)#
neighbor {ip-address | peer-group-name} advertisement-interval seconds
• Changes the default time interval in the sending of BGP routing updates for a specific neighbor:
– If lowered, can improve convergence
– Can consume considerable resources in a jittery network if the value is set too low
• Default values:
– 30 seconds for EBGP neighbors
– 5 seconds for IBGP neighbors
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-16

With the help of the **neighbor advertisement-interval** command, network administrators can modify the default advertisement interval for a specific BGP peer. BGP advertisements are rate-limited by the router using the advertisement interval timer (defined as the parameter MinRouteAdvertisementInterval in RFC 1771). When a BGP-speaking router sends a route update to a neighbor for a specific destination, it is not allowed to send another update to the neighbor about the same destination until a period of time equal to the advertisement interval has elapsed. In that way, the advertisement interval timer acts as a form of rate limiting on a per-destination basis, even though the value of the advertisement interval is configured for each neighbor.

It is important to note that during the time that the router is waiting for the advertisement interval timer to expire, the router can still receive and process route updates from BGP neighbors. Using the **neighbor advertisement-interval** command does not rate-limit BGP route selection (inbound updates and subsequent processing), but only the rate of outgoing route advertisements.

When faster propagation of successive BGP updates (which are batched and rate-limited) is required, the network administrator can lower the default value of the advertisement interval, thus improving convergence.

In routers dealing with large BGP tables and/or less stable networks, lowering the advertisement interval could potentially lead to consuming large portions of the router resources.

neighbor advertisement-interval

To set the minimum interval in the sending of BGP routing updates, use the **neighbor advertisement-interval** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **advertisement-interval** *seconds*

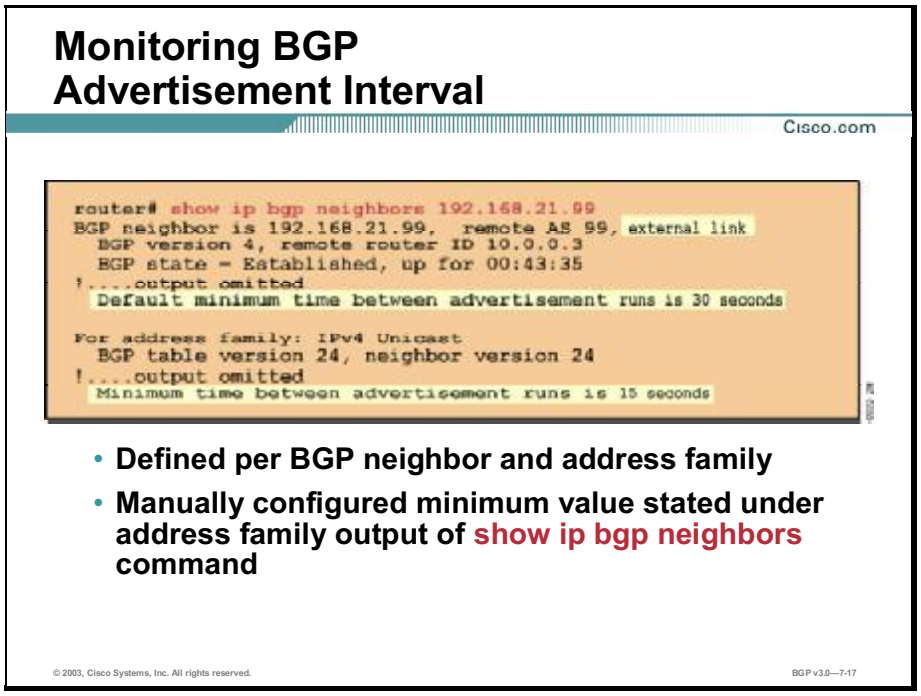
To remove an entry, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **advertisement-interval** *seconds*

Syntax Description

ip-address	Neighbor IP address.
peer-group-name	Name of a BGP peer group. If a BGP peer group is specified by using the *peer-group-name* argument, all members of the peer group will inherit the characteristic configured with this command.
seconds	Time in seconds. Integer from 0 to 600. The default is 30 seconds for external peers and 5 seconds for internal peers.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring BGP Advertisement Interval
Cisco.com
router# show ip bgp neighbors 192.168.21.99
BGP neighbor is 192.168.21.99, remote AS 99, external link
BGP version 4, remote router ID 10.0.0.3
BGP state = Established, up for 00:43:35
!....output omitted
Default minimum time between advertisement runs is 30 seconds
For address family: IPv4 Unicast
BGP table version 24, neighbor version 24
!....output omitted
Minimum time between advertisement runs is 15 seconds
• Defined per BGP neighbor and address family
• Manually configured minimum value stated under address family output of show ip bgp neighbors command
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-17

You can examine the currently configured BGP advertisement interval with the **show ip bgp neighbors** command. The advertisement interval is defined per a specific neighbor belonging to a specific BGP address family. Actual values of the advertisement interval are therefore stated under the specific address family portion of the neighbor output.

Practice

Q1) How does the advertisement interval timer act as a rate-limiting feature in BGP?

A) The advertisement interval timer specifies the time between successive runs of the BGP scanner process, resulting in rate-limiting the update process.

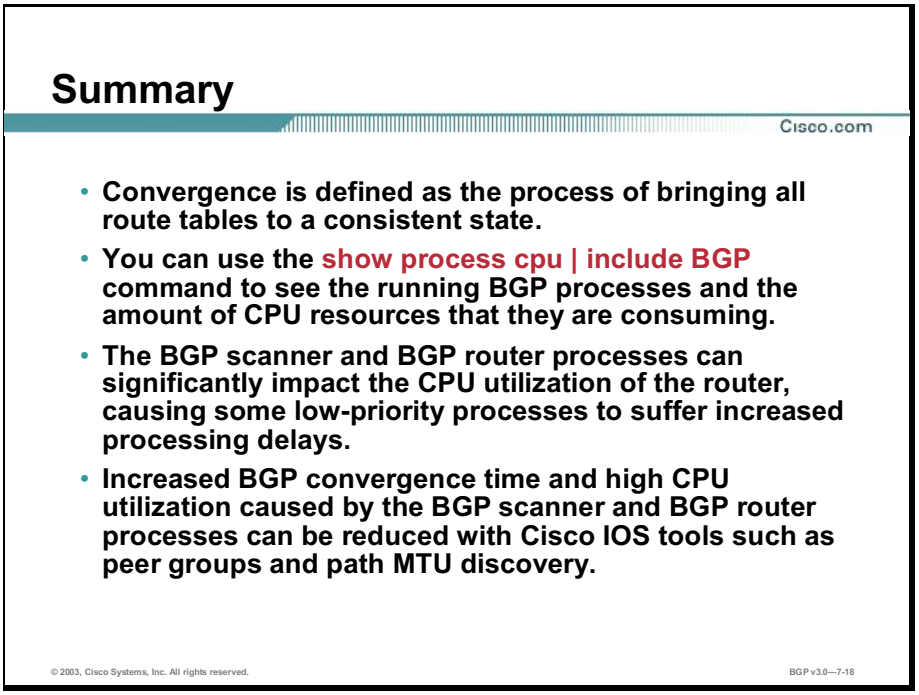
B) Adjusting the advertisement interval limits the rate at which the router sends and receives route advertisements.

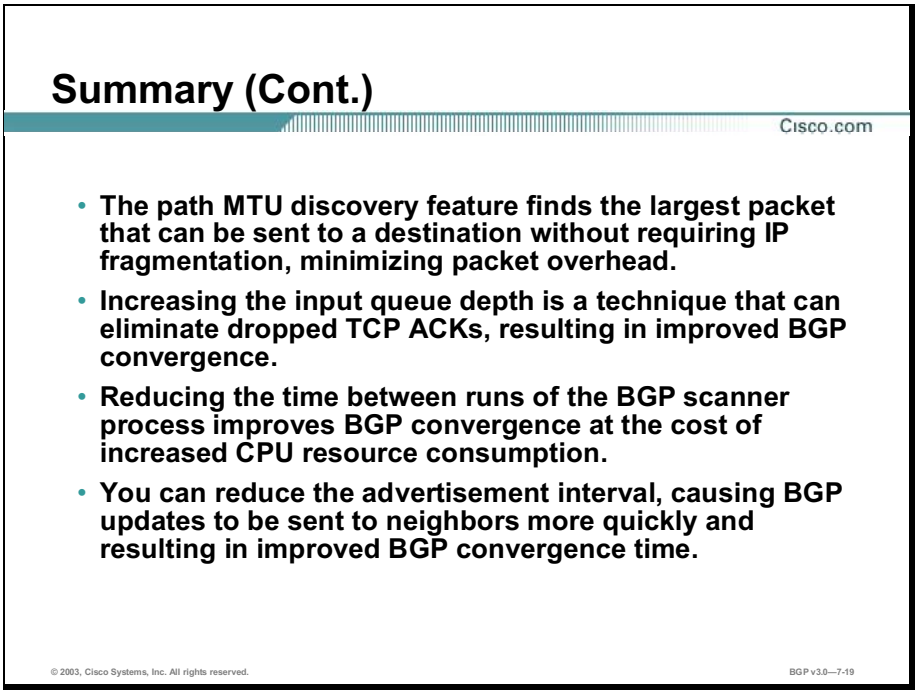
C) Adjusting the advertisement interval does not limit the rate of route selection but only the rate of route advertisement.

D) Adjusting the advertisement interval limits the rate at which the router can receive BGP updates from its neighbors.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Convergence is defined as the process of bringing all route tables to a consistent state.
• You can use the show process cpu | include BGP command to see the running BGP processes and the amount of CPU resources that they are consuming.
• The BGP scanner and BGP router processes can significantly impact the CPU utilization of the router, causing some low-priority processes to suffer increased processing delays.
• Increased BGP convergence time and high CPU utilization caused by the BGP scanner and BGP router processes can be reduced with Cisco IOS tools such as peer groups and path MTU discovery.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-18


Summary (Cont.)
Cisco.com
• The path MTU discovery feature finds the largest packet that can be sent to a destination without requiring IP fragmentation, minimizing packet overhead.
• Increasing the input queue depth is a technique that can eliminate dropped TCP ACKs, resulting in improved BGP convergence.
• Reducing the time between runs of the BGP scanner process improves BGP convergence at the cost of increased CPU resource consumption.
• You can reduce the advertisement interval, causing BGP updates to be sent to neighbors more quickly and resulting in improved BGP convergence time.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-19

 Copyright © 2003, Cisco Systems, Inc.

Next Steps

After completing this lesson, go to:

- Limiting the Number of Prefixes Received from a BGP Neighbor lesson

References

For additional information, refer to these resources:

- For more information on the Cisco IOS commands presented in this lesson, refer to "Cisco IOS IP and IP Routing Command Reference, Release 12.1" at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121cgcr/ip_r/index.htm

- For more information on BGP convergence, refer to "An Analysis of BGP Convergence Properties," by T. Griffin and G. T. Wilfong, SIGCOMM 1999

- For further information on BGP convergence, refer to "Improving BGP Convergence Through Consistency Assertions," by Dan Pei, Xiaoliang Zhao, Lan Wang, Dan Massey, Allison Mankin, S. Felix Wu, and Lixia Zhang, 2002

Quiz: Improving BGP Convergence

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe convergence in BGP networks
- Describe the function of BGP router processes
- Describe the effects of BGP processes on router CPU resources
- Identify Cisco IOS performance improvements to reduce BGP convergence time
- Identify the commands to configure and monitor path MTU discovery
- Identify the commands to configure and monitor the input queue depth on a router interface
- Identify the commands to configure and monitor BGP scan time
- Identify the commands to configure and monitor the BGP advertisement interval

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What are three characteristics of a converged BGP network? (Choose three.)

A) The input queue and output queue for all peers is 0.

B) All routes in the BGP table have been installed in the routing table.

C) The table version for all peers equals the table version of the BGP table.

D) All routes have been accepted.

 Copyright © 2003, Cisco Systems, Inc.

Q2) Which two of the following modifications result in improved BGP convergence? (Choose two.)

A) increasing the default value of BGP hold time

B) lowering the default value of BGP scan time

C) increasing the default value of the neighbor advertisement intervals

D) lowering the default value of the neighbor advertisement intervals

Q3) What is the main task of the BGP scanner process?

A) send routing updates to BGP neighbors

B) walk the BGP table for routes to enter into the IP routing table

C) confirm the reachability of BGP next hops

D) scan the router configuration to establish and maintain BGP neighbors

Q4) One of your BGP core routers is experiencing periodic slow responses to ping packets directed to it from the network management console. The router has just been configured to receive full Internet routes, and you suspect that the BGP router process is causing CPU utilization issues in the core router. Which two router commands should you use to confirm your suspicion? (Choose two.)

A) **show ip route**

B) **show ip bgp summary**

C) **show process cpu**

D) **show memory**

Q5) The output of a **show interfaces fastethernet 0/0** command follows:

```
Fast Ethernet0 is up, line protocol is up
  Hardware is DEC21140, address is 0000.0c0c.1111 (bia
0002.eaa3.5a60)
  Internet address is 112.64.101.17 255.255.255.240
  MTU 1460 bytes, BW 100000 Kbit, DLY 100 usec, rely 255/255, load
200/255
  Encapsulation ARPA, loopback not set, keepalive not set, hdx,
100BaseTX
  ARP type: ARPA, ARP Timeout 4:00:00
  Last input never, output 0:00:16, output hang 0:28:01
  Last clearing of "show interface" counters 0:20:05
  Output queue 25/40, 0 drops; input queue 50/500, 1470 drops
  5 minute input rate 21666400 bits/sec, 1855 packets/sec
  5 minute output rate 72221 bits/sec, 618 packets/sec
```

Q6) How has the interface been modified to improve BGP convergence?

A) The output queue has been decreased to expedite packet forwarding out the fast Ethernet interface.

B) The drop threshold of the input queue has been set to begin randomly discarding packets after the queue reaches 50 packets deep.

C) PMTU discovery has been enabled, setting the interface MSS to 1460 bytes.

D) The size of the input queue has been increased to support up to 500 incoming packets.

Q7) Refer to the following Cisco IOS router output:

```
router# show ip bgp summary
BGP router identifier 172.16.0.4, local AS number 1
BGP table version is 16, main routing table version 16
20 network entries and 20 paths using 2826 bytes of memory
8 BGP path attribute entries using 480 bytes of memory
7 BGP AS-PATH entries using 168 bytes of memory
3 BGP community entries using 72 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
3 BGP filter-list cache entries using 36 bytes of memory
BGP activity 20/0 prefixes, 24/4 paths, scan interval 120 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
172.16.0.1      4     1      30      30       16    0    0 00:23:13
5
172.16.0.2      4     1      33      30       16    0    0 00:23:15
5
172.16.0.3      4     1      27      30       16    0    0 00:23:14
5
192.168.21.99   4    99      31      35       16    0    0 00:23:04
5
```

What two parameters would indicate that the BGP network has converged? (Choose two.)

A) The TblVer for all neighbors is 16.

B) V is set to 4 for all neighbors.

C) The InQ and OutQ for all neighbors is 0.

D) All neighbors are in the Established state and have the same PfxRcd value.

 Copyright © 2003, Cisco Systems, Inc.

Q8) Using the command output from question 6, identify how frequently the BGP scanner process will run on the router.

A) By default, the process will run every 60 seconds.

B) The process has run 16 times and will run again when the next BGP update arrives.

C) The process will run on this router every 120 seconds.

D) It cannot be determined from this output.

Q9) What are two potential issues caused by modifying the default scan time and advertisement interval on a BGP router? (Choose two.)

A) Router CPU resources can be exhausted.

B) Router memory resources can be depleted.

C) Routing loops are more likely.

D) BGP could converge faster than the IGP, causing network black holes.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

 Copyright © 2003, Cisco Systems, Inc.

Limiting the Number of Prefixes Received from a BGP Neighbor

Overview

The Border Gateway Protocol (BGP) is designed for reliability and scalability. As such, BGP has a tremendous amount of flexibility regarding administrative policy controls, route selection, and performance tuning and scalability features. This lesson introduces an advanced BGP configuration tool designed to improve BGP scalability and performance by reducing the number of prefixes that a router receives from a BGP neighbor. Discussed in this lesson is the need for prefix limiting and how to configure and monitor the maximum-prefix function.

Importance

There are currently more than 110,000 prefixes on the Internet today. There are many circumstances where network administrators do not need or desire their routers to carry full Internet routing. Furthermore, there is a need to provide protective controls on customer-facing routers to ensure that a configuration error does not cause the accidental advertisement of prefixes from autonomous systems that did not originate them.

Objectives

Upon completing this lesson, you will be able to:

- Describe the need for limiting the number of routes received from a BGP neighbor
- Identify the Cisco IOS® commands required to configure the BGP maximum-prefix function
- Identify the Cisco IOS commands required to monitor the BGP maximum-prefix function

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module

Outline

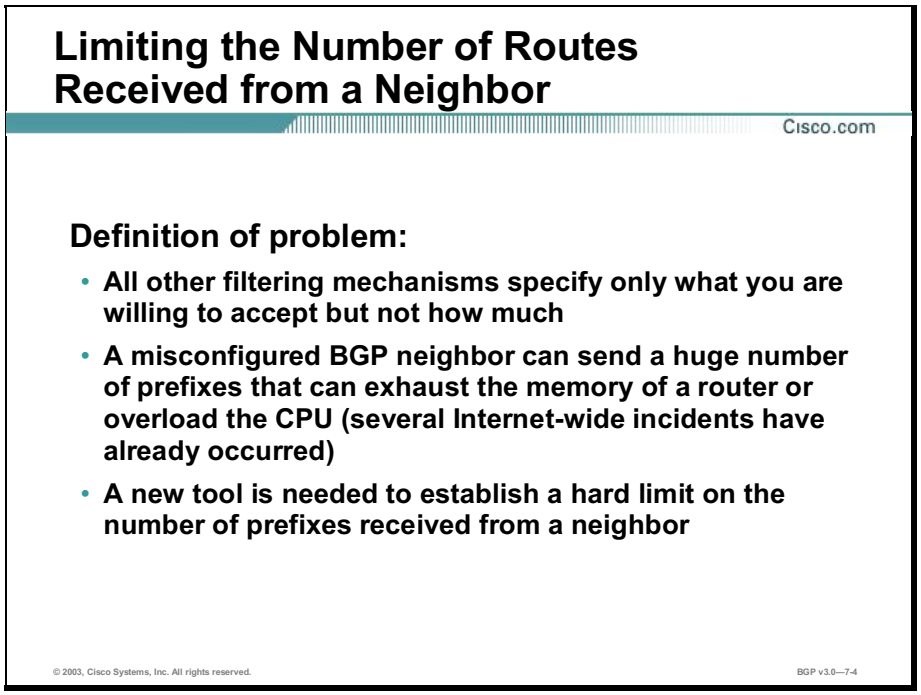
This lesson includes these topics:

- Overview
- Limiting the Number of Routes Received from a Neighbor
- Configuring the BGP Maximum-Prefix Function
- Monitoring the BGP Maximum-Prefix Function
- Summary
- Assessment (Lab): Limiting the Number of Prefixes Received from a BGP Neighbor

 Copyright © 2003, Cisco Systems, Inc.

Limiting the Number of Routes Received from a Neighbor

This topic describes the need for limiting the number of routes (prefixes) received from a BGP neighbor.


Limiting the Number of Routes Received from a Neighbor
Cisco.com
Definition of problem:
• All other filtering mechanisms specify only what you are willing to accept but not how much
• A misconfigured BGP neighbor can send a huge number of prefixes that can exhaust the memory of a router or overload the CPU (several Internet-wide incidents have already occurred)
• A new tool is needed to establish a hard limit on the number of prefixes received from a neighbor
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-4

Incoming filters and route-maps indicate what BGP attribute values that a route should have in order to be accepted. Route filters can be applied that match routes based on the network number or the BGP attributes attached to a route. The most commonly applied filter is one that matches the contents of the AS-path attribute.

An Internet service provider (ISP) with a multihomed customer may use filters to ensure that the routes received from the customer originate within the autonomous system (AS) of the customer. Using an AS-path access-list is one method of achieving this goal.

An improperly configured filter in a customer router may accidentally cause a large number of Internet routes to be received by the customer. Even worse, a faulty configuration may cause prefixes to be advertised by the customer as if the routes originated inside the customer AS. This situation would result in a BGP table, in the ISP router, that lists many of the possible destination networks on the Internet as reachable in the customer AS (with the AS path containing only a single entry, the customer AS). The BGP route selection in the ISP network would select those routes as the best (based on the AS-path length) and direct much of the provider traffic intended for the Internet to the customer network.

An AS-path filter in the ISP router would not prevent this accident. The routes sent by the customer have the anticipated AS-path value. A prefix-list that distinctly identifies and permits each of the network numbers that the customer may advertise would have prevented the accident. But such a prefix-list is hard to maintain.

A more scalable solution is to use a new tool, which limits the number of routes that are received from a specific neighbor.

Practice

Q1) What are three reasons to limit the number of BGP prefixes received from a neighbor? (Choose three.)

A) to prevent denial-of-service attacks

B) to protect against incorrect router configuration on the neighbor side

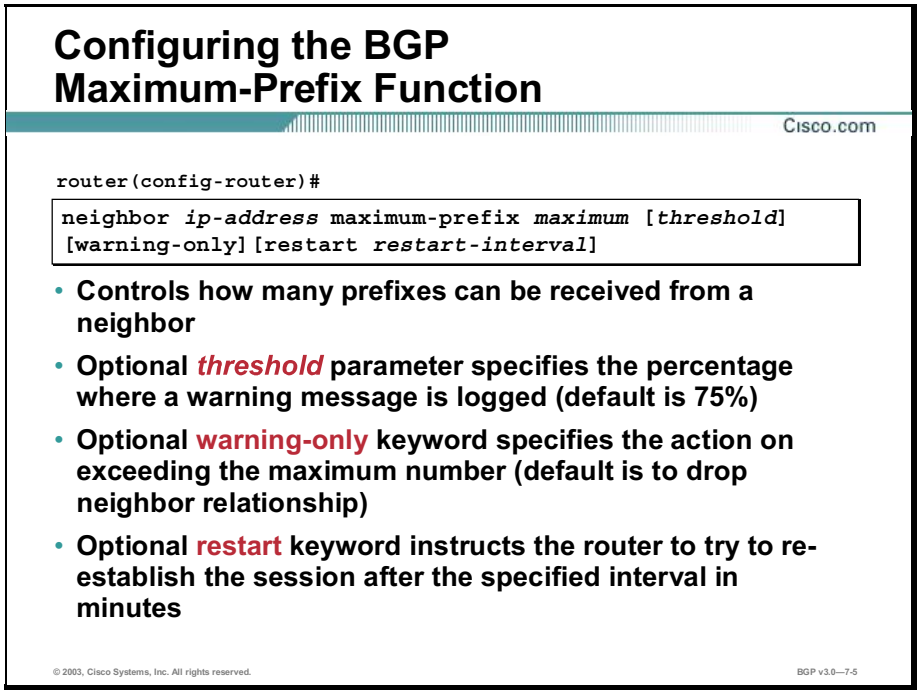
C) to prevent redundant routing information from being loaded into the BGP table

D) to avoid overloading router memory and CPU resources

 Copyright © 2003, Cisco Systems, Inc.

Configuring the BGP Maximum-Prefix Function

This topic presents the Cisco IOS command used to configure the maximum-prefix function in BGP.


Configuring the BGP Maximum-Prefix Function
Cisco.com
router(config-router)#
neighbor ip-address maximum-prefix maximum [threshold] [warning-only] [restart restart-interval]
• Controls how many prefixes can be received from a neighbor
• Optional threshold parameter specifies the percentage where a warning message is logged (default is 75%)
• Optional warning-only keyword specifies the action on exceeding the maximum number (default is to drop neighbor relationship)
• Optional restart keyword instructs the router to try to re-establish the session after the specified interval in minutes
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-5

neighbor maximum-prefix

To control how many prefixes that a BGP router can receive from a neighbor, use the **neighbor maximum-prefix** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **maximum-prefix** *maximum* [*threshold*] [**warning-only**] [**restart** *restart-interval*]

To disable this function, use the **no** form of this command.

no neighbor {*ip-address* | *peer-group-name*} **maximum-prefix** *maximum*

Syntax Description

ip-address	IP address of the neighbor.
peer-group-name	Name of a BGP peer group.
maximum	Maximum number of prefixes allowed from this neighbor.
threshold	(Optional) Integer specifying at what percentage of maximum the router starts to generate a warning message. The range is 1 to 100 percent; the default is 75 percent.

warning-only	(Optional) Allows the router to generate a log message when the maximum is exceeded, instead of terminating the peering.
restart	(Optional) Configures the router to automatically re-establish a peering session that has been disabled because the maximum-prefix limit has been exceeded. The configurable range of the restart interval is from 1 to 65,535 minutes.

This command allows you to configure a maximum number of prefixes that a BGP router is allowed to receive from a peer. It adds another mechanism (in addition to distribute-lists, filter-lists, and route-maps) to control prefixes received from a peer.

When the number of received prefixes exceeds the maximum number configured, the router terminates the peering (by default). However, if the **warning-only** keyword is configured, the router sends only a log message but continues peering with the sender. If the peer is terminated, the peer session will remain down until the **clear ip bgp** command is issued on the router, unless you have included the **restart** keyword in the configuration.

Note You can use the **bgp dampening** command to configure the dampening of a flapping route or interface when a peer is sending too many prefixes and causing network instability. You should need the **restart** command only when you are troubleshooting or tuning a router that is sending an excessive number of prefixes.

Practice

Q1) By default, what happens when the number of configured prefixes is exceeded?

A) A filter is automatically generated and applied to incoming BGP updates.

B) A syslog message is sent to the log.

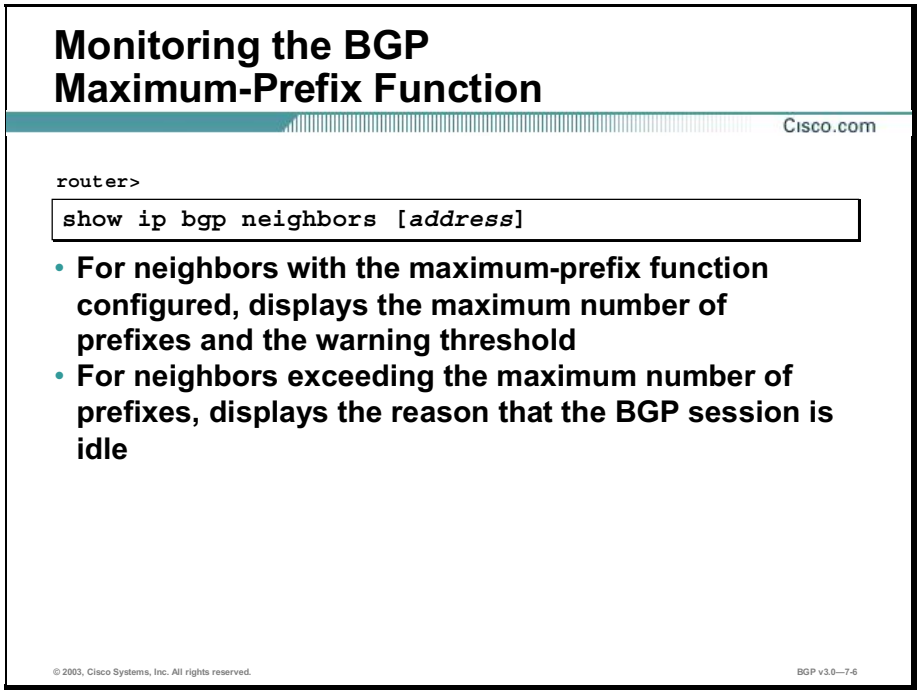
C) The BGP session is restarted to clear out the BGP table.

D) The BGP session is terminated and will remain inactive until the session is manually enabled once again.

 Copyright © 2003, Cisco Systems, Inc.

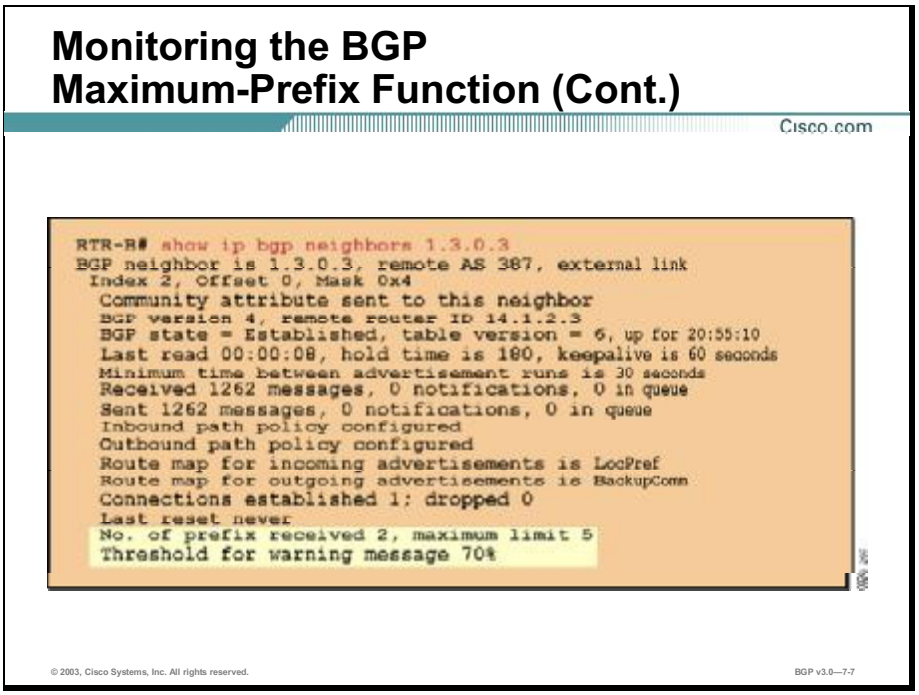
Monitoring the BGP Maximum-Prefix Function

This topic presents the Cisco IOS command used to monitor the operation of the maximum-prefix function in BGP.


Monitoring the BGP Maximum-Prefix Function
Cisco.com
router>
show ip bgp neighbors [address]
• For neighbors with the maximum-prefix function configured, displays the maximum number of prefixes and the warning threshold
• For neighbors exceeding the maximum number of prefixes, displays the reason that the BGP session is idle
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-6

Network administrators use he **show ip bgp neighbors** command to monitor the status of BGP neighbors. Among other things, the command displays information about how many prefixes that a BGP router has received from a neighbor and if any limits have been configured.

Example


Monitoring the BGP
Maximum-Prefix Function (Cont.)
Cisco.com
RTR-R# show ip bgp neighbors 1.3.0.3
BGP neighbor is 1.3.0.3, remote AS 387, external link
Index 2, Offset 0, Mask 0x4
Community attribute sent to this neighbor
BGP version 4, remote router ID 14.1.2.3
BGP state = Established, table version = 6, up for 20:55:10
Last read 00:00:08, hold time is 180, keepalive is 60 seconds
Minimum time between advertisement runs is 30 seconds
Received 1262 messages, 0 notifications, 0 in queue
Sent 1262 messages, 0 notifications, 0 in queue
Inbound path policy configured
Outbound path policy configured
Route map for incoming advertisements is LocPref
Route map for outgoing advertisements is BackupComm
Connections established 1; dropped 0
Last reset never
No. of prefix received 2, maximum limit 5
Threshold for warning message 70%
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-7

In this example, the neighbor with IP address 1.3.0.3 has been configured with a prefix limit of five prefixes. Currently, the BGP router has received two prefixes, which is under the limit. The threshold for the warning message is set to 70 percent, meaning that after receiving four prefixes from the 1.3.0.3 neighbor, the BGP router will generate a warning message.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring the BGP Maximum-Prefix Function (Cont.)

Cisco.com

```
RTR-B#
%BGP-4-MAXPFX: No. of prefix received from 1.3.0.3 reaches 4, max 5
%BGP-3-MAXPFXEXCEED: No. of prefix received from 1.3.0.3: 6 exceed limit 5

RTR-B# show ip bgp sum
BGP table version is 22, main routing table version 22
9 network entries (9/27 paths) using 1920 bytes of memory
5 BGP path attribute entries using 572 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
1.0.0.1         4   213    1269    1268       22    0    0 21:02:19        0
1.3.0.3         4   387    1272    1274        0    0    0 00:00:08 Idle

RTR-B# show ip bgp neighbor 1.3.0.3
BGP neighbor is 1.3.0.3, remote AS 387, external link
 ...
  Last reset 00:00:10, due to : Peer exceeding maximum prefix limit
  Peer had exceeded the max. no. of prefixes configured.
  Reduce the no. of prefix and clear ip bgp 1.3.0.3 to restore peering
  No active TCP connection
```

© 2003, Cisco Systems, Inc. All rights reserved.

The logging outputs, in the example above, show that a BGP neighbor is close to exceeding the configured maximum-prefix limit. The total number of received prefixes has reached four, which is over the threshold to generate a warning message. The warning is displayed on the console and optionally sent to a syslog server.

The logging output now shows that two more prefixes have been received. The total number is now six, which is above the configured limit. The BGP session is therefore terminated. The output of the **show ip bgp neighbors** command shows that the reason for resetting the session is that the peer exceeded the configured maximum number of prefixes. As a result of the session being reset, the BGP session will remain in the Idle state.

To force the neighbor from the Idle state into the Active state and to re-establish the BGP session, the network administrator must issue the **clear ip bgp ip-address** command for the neighbor (except if the network administrator has specified the **restart** keyword in the configuration, in which case the router tries to re-establish the BGP session automatically after the expiration of the configured restart timeout interval).

Practice

Q1) In what two situations would a directly connected BGP neighbor stay in the Idle state? (Choose two.)

A) The neighbor has exceeded the maximum number of allowed prefixes.

B) The maximum-prefix threshold has been reached.

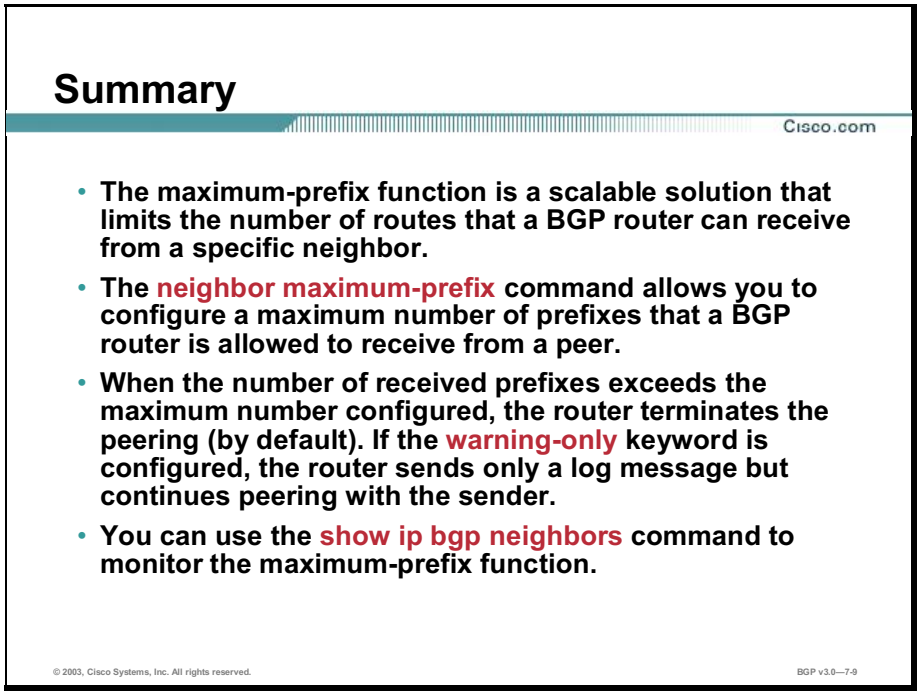
C) The **restart** option has not been specified with the **maximum-prefix** command.

D) The neighbor is more than one hop away.

 Copyright © 2003, Cisco Systems, Inc.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• The maximum-prefix function is a scalable solution that limits the number of routes that a BGP router can receive from a specific neighbor.
• The neighbor maximum-prefix command allows you to configure a maximum number of prefixes that a BGP router is allowed to receive from a peer.
• When the number of received prefixes exceeds the maximum number configured, the router terminates the peering (by default). If the warning-only keyword is configured, the router sends only a log message but continues peering with the sender.
• You can use the show ip bgp neighbors command to monitor the maximum-prefix function.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-9

Next Steps

After completing this lesson, go to:

- BGP Peer Groups lesson

References

For additional information, refer to these resources:

- For more information on the **neighbor maximum-prefix** command and its function, refer to "BGP Restart Session After Maximum-Prefix Limit" at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120limit/120s/120s22/sbgprsmp.htm

Laboratory Exercise: Limiting the Number of Prefixes Received from a BGP Neighbor

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to limit the number of prefixes received from a neighbor.

After completing this exercise, you will be able to:

- Configure the BGP maximum-prefix function
- Monitor the BGP maximum-prefix function

 Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor *ip-address* **maximum-prefix** *maximum* [**threshold**] [**warning-only**][**restart** *restart-interval*]	Specify the allowed maximum number of prefixes.
show ip bgp	Inspect the contents of the BGP table.
show ip neighbor *neighbor*	Show detailed information about a specific neighbor.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- You have noticed that one of your upstream service providers occasionally announces a very large number of BGP prefixes, resulting in memory and CPU overload on your router (WGxR1). You suspect that the service provider might have internal problems, and you would like to be protected against these problems. The service provider is asking for evidence that you are receiving more prefixes than you should.

- In this exercise, you will configure logging on your router to warn you if you receive more BGP prefixes than you should from your service provider.

- Your analysis shows that you must perform the following tasks to minimize the size of the BGP table on router WGxR1:

  – Create a warning log whenever your service provider (“Cheap”) sends you more than 15 prefixes.

- Figure 1 shows the connectivity that is established between your AS and the two service providers “Good” and “Cheap.”

Copyright © 2003, Cisco Systems, Inc.

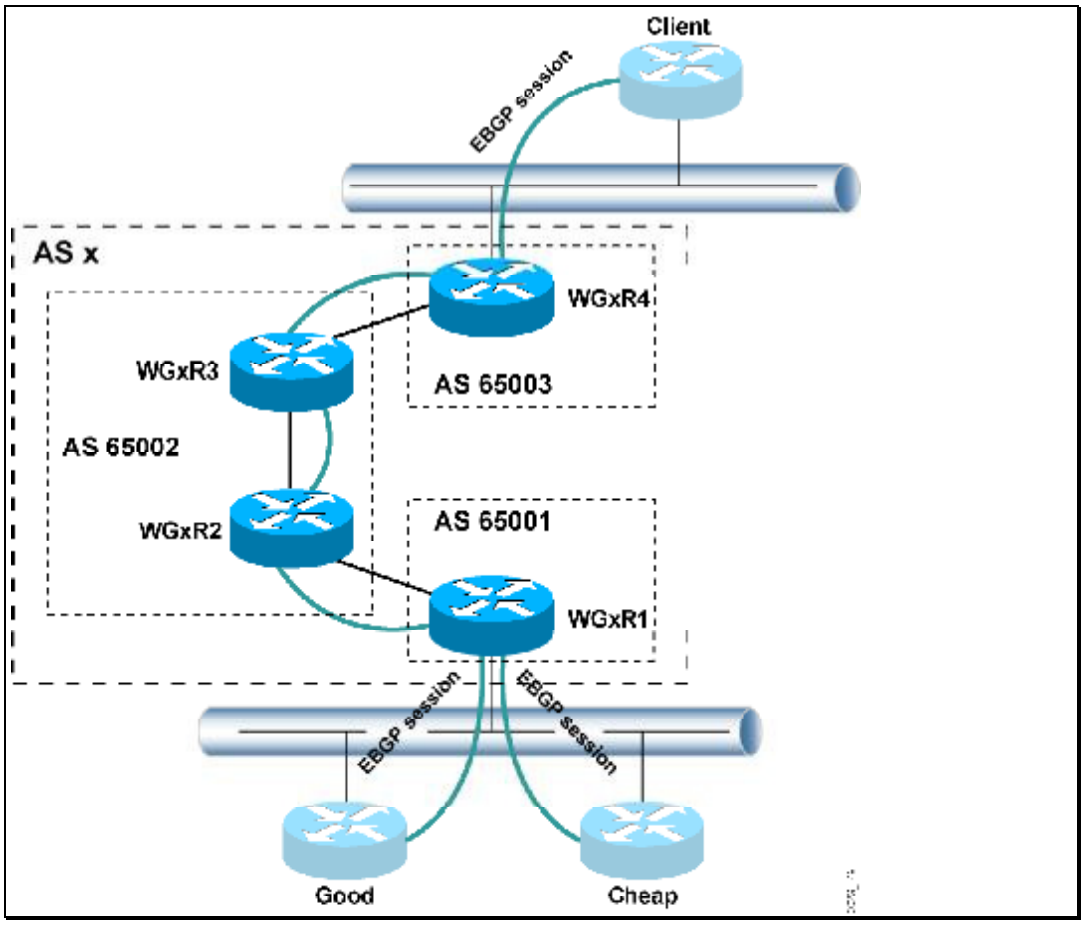


Figure 1: Connections to two different service providers

Exercise Procedure

Complete these steps:

Configuring the maximum-prefix function

Step 1 Set a limit to the number of prefixes that your router WGxR1 accepts from router "Cheap." Do not disable BGP, but instead create a warning message in the router log.

Step 2 Check your console for the following two messages:

```
22:46:40: %BGP-4-MAXPFX: No. of prefix received from 192.168.20.22
(afi 0) reaches 12, max 15
22:49:00: %BGP-3-MAXPFXEXCEED: No. of prefix received from
192.168.20.22 (afi 0): 16 exceed limit 15
```

Step 3 After showing the service provider your logs, you have decided to take administrative action by implementing a BGP policy that will terminate the neighbor session between you and your upstream provider ("Cheap") should it continue to violate your agreement on the number of prefixes advertised to your network. Configure your router such that the BGP session is terminated if your router (WGxR1) receives more than the agreed number of prefixes.

 Copyright © 2003, Cisco Systems, Inc.

Exercise Verification

You have completed this exercise when you attain these results:

- Verify if both neighbors are up. If not, find out why.

```
BGP-4-MAXPFX: No. of prefix received from 192.168.20.22 (afi 0)
reaches 12, max 15
BGP-3-MAXPFXEXCEED: No. of prefix received from 192.168.20.22 (afi
0): 16 exceed limit 15
BGP-5-ADJCHANGE: neighbor 192.168.20.22 Down BGP Notification sent
BGP-3-NOTIFICATION: sent to neighbor 192.168.20.22 3/1 (update
malformed) 0 bytes

WG1R1#sh ip bgp summary
BGP router identifier 197.1.8.1, local AS number 1
BGP table version is 72, main routing table version 72
41 network entries and 41 paths using 5453 bytes of memory
18 BGP path attribute entries using 936 bytes of memory
10 BGP AS-PATH entries using 256 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
5 BGP filter-list cache entries using 60 bytes of memory
16 received paths for inbound soft reconfiguration
BGP activity 300/2803 prefixes, 522/469 paths, scan interval 15 secs

Neighbor        V   AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.20.20   4   20    1504    1451       72    0    0 00:26:44
10
192.168.20.22   4   22    1123    1073        0    0    0 00:00:37
Idle (PfxCt)

WG1R1#sh ip bgp neighbor 192.168.20.22
BGP neighbor is 192.168.20.22,  remote AS 22, external link
  BGP version 4, remote router ID 0.0.0.0
  BGP state = Idle
  Last read 00:04:48, hold time is 180, keepalive interval is 60
seconds
  Received 1123 messages, 0 notifications, 0 in queue
  Sent 1073 messages, 1 notifications, 0 in queue
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 30 seconds
  Default weight 200

 For address family: IPv4 Unicast
  BGP table version 72, neighbor version 0
  Index 2, Offset 0, Mask 0x4
  Inbound soft reconfiguration allowed
  Inbound path policy configured
  Outbound path policy configured
  Incoming update prefix filter list is FromCheap
  Incoming update AS path filter list is 2
  Outgoing update AS path filter list is 3, maximum limit 15
```

```
  Threshold for warning message 75%

  Connections established 6; dropped 6
  Last reset 00:04:48, due to Error during connection collision
  Peer had exceeded the max. no. of prefixes configured.
  Reduce the no. of prefix and clear ip bgp 192.168.20.22 to restore
peering
  No active TCP connection
WG1R1#
```

- Increase the number of allowed prefixes to continue testing. Clear the neighbor session between your WGxR1 and router "Cheap" afterward.

- Inspect the prefixes that your router is receiving from "Cheap" with the **show ip bgp neighbors** command. Be sure to use the **routes** keyword.

```
WG1R1#sh ip bgp nei 192.168.20.22 routes
BGP table version is 249, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? – incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.0.0.0         192.168.20.233                         0 22 20 i
*  128.20.0.0       192.168.20.22            0             0 22 i
*  128.20.12.0/24   192.168.20.20                          0 22 20 i
*  128.22.0.0       192.168.20.22            0             0 22 i
*> 128.22.12.0/24   192.168.20.22            0             0 22 i
*> 128.26.0.0       192.168.20.22            0             0 22 26 i
*  128.37.0.0       192.168.20.22            0             0 22 26 42
37 i
*  128.42.0.0       192.168.20.22            0             0 22 26 42
i
*> 128.51.0.0       192.168.20.22            0             0 22 26 51
i
*  128.213.0.0      192.168.20.20                          0 22 20
213 i
*> 128.214.0.0      192.168.20.22            0             0 22 214 i
*  192.20.11.0      192.168.20.22            0             0 22 i
*  192.20.12.0/30   192.168.20.20                          0 22 20 i
*  192.22.11.0      192.168.20.22            0             0 22 i
*> 192.22.12.0/30   192.168.20.22            0             0 22 i
*> 192.26.11.0      192.168.20.22            0             0 22 26 i
*  192.37.11.0      192.168.20.22            0             0 22 26 42
37 i
*  192.42.11.0      192.168.20.22            0             0 22 26 42
i
*> 192.51.11.0      192.168.20.22            0             0 22 26 51
i
*  192.213.11.0     192.168.20.20                          0 22 20
213 i
*> 192.214.11.0     192.168.20.22            0             0 22 214 i
```

 Copyright © 2003, Cisco Systems, Inc.

```
*  200.20.0.0/16    192.168.20.20                          0 22 20 i
*> 200.22.0.0/16    192.168.20.22           0              0 22 i

Total number of prefixes 23
```

- Increase the number of allowed prefixes so that all of the prefixes coming from “Cheap” will be allowed and the session will not be terminated.

- Clear the neighbor session (configuration done in this exercise) between your WGxR1 and router “Cheap” afterward and ensure that the neighbor session remains active.

Answer this question:

Q1) What can you configure on your router to allow the BGP session to restart automatically if the received number of prefixes exceeds the configured maximum?

 Copyright © 2003, Cisco Systems, Inc.

BGP Peer Groups

Overview

Typical service provider networks usually contain Border Gateway Protocol- (BGP-) speaking routers consisting of many neighbors configured with the same administrative policies (such as outbound route-maps, distribute-lists, filter-lists, update source, and so on). Network administrators can group together neighbors with the same update policies into peer groups to simplify configuration and, more importantly, to make BGP updates more efficient.

As a result of using peer groups, the same update packets can be sent from a router to each group member, resulting in a reduction in the number of CPU cycles that BGP requires to advertise routes to peers. This lesson introduces peer groups as a BGP scalability mechanism. The lesson also discusses the commands required to properly configure and monitor BGP peer groups.

Importance

Scaling routers to meet the demands of full Internet routing and associated administrative policies requires protocols like BGP with embedded scalability mechanisms. In environments where network administrators must configure a large number of BGP peers, peer groups are a scalability tool that reduces both administrative overhead and router resource requirements. The major benefit of specifying a BGP peer group is that it reduces the amount of system resources (CPU and memory) used in the generation of an update. It also simplifies BGP configuration because it allows the routing table to be checked only once and updates to be replicated to all other in-sync peer group members.

Objectives

Upon completing this lesson, you will be able to:

- Describe the need for BGP peer groups
- Describe the performance benefits of using BGP peer groups

- Describe the limitations of BGP peer groups
- Describe the implementation of BGP peer groups in Cisco IOS® software
- Identify the Cisco IOS commands required to configure BGP peer groups
- Identify the Cisco IOS commands required to monitor the operation of BGP peer groups

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

Outline

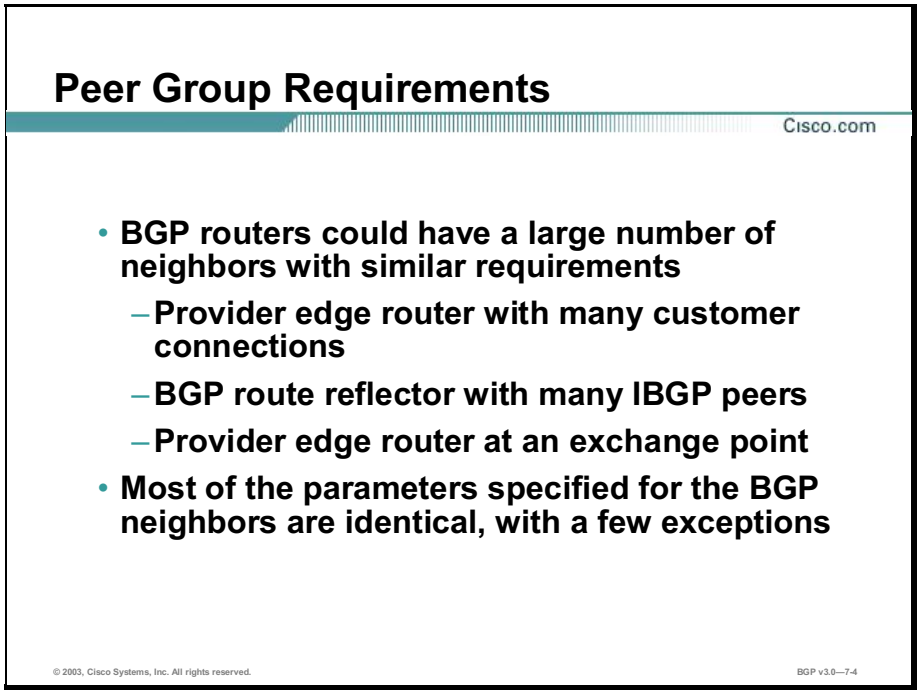
This lesson includes these topics:

- Overview
- Peer Group Requirements
- Peer Groups as a BGP Performance Tool
- BGP Peer Group Limitations
- BGP Peer Groups in Cisco IOS Software
- Configuring Peer Groups
- Monitoring Peer Groups
- Summary
- Assessment (Lab): BGP Peer Groups

 Copyright © 2003, Cisco Systems, Inc.

Peer Group Requirements

This topic describes the need for BGP peer groups and provides three examples where network administrators can use peer groups in service provider networks.

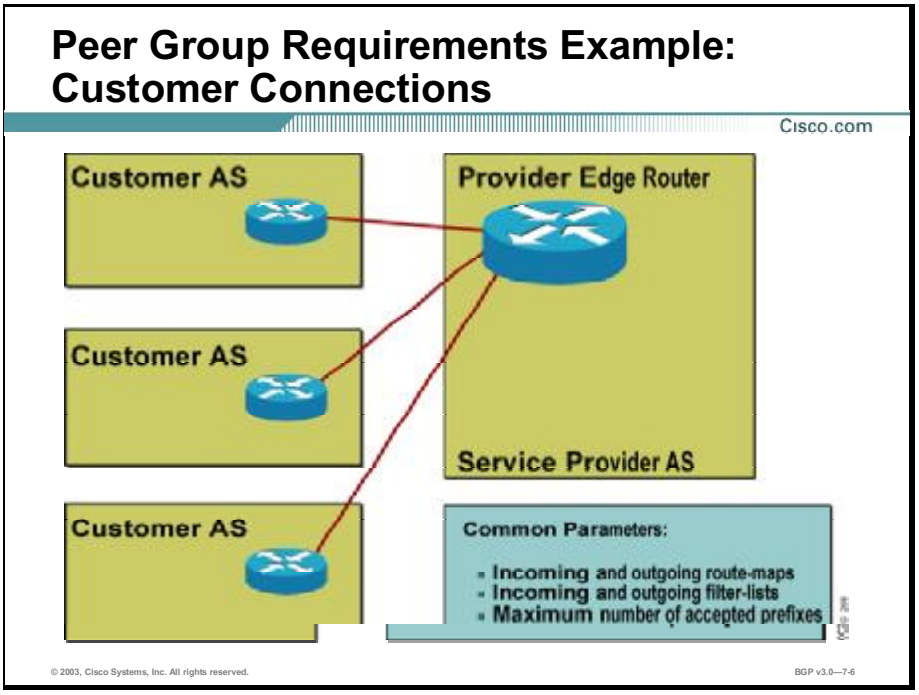

Peer Group Requirements
Cisco.com
• BGP routers could have a large number of neighbors with similar requirements
– Provider edge router with many customer connections
– BGP route reflector with many IBGP peers
– Provider edge router at an exchange point
• Most of the parameters specified for the BGP neighbors are identical, with a few exceptions
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-4

In many cases, a network administrator must configure a single router with a large number of neighbors, each neighbor having similar parameters as the others. This situation may cause time-consuming configuration work, because the network administrator has to configure almost the same filter-list, route-map, and prefix-list for all of the neighbors. A service provider network with an edge router having a large number of customers attached to it, where each customer requires its own BGP session, may find that all of the BGP sessions to its customer routers have almost identical configurations.

Likewise, Internal Border Gateway Protocol (IBGP) sessions are almost always identically configured. If a full mesh is deployed within an autonomous system (AS), a large number of peer configurations might exist. Recall that an AS containing only 15 routers will require ([5 * 14]/ 2) = 105 neighbor sessions to meet the full-mesh requirement of BGP. Configuring 105 neighbors with duplicate parameters leads to a tremendous amount of redundant configuration.

To ease the burden of configuring a large number of neighbors with identical or similar parameters (for example, route-maps, filter-lists, or prefix-lists), the concept of peer groups was introduced. The network administrator can configure a template, called the peer group. The administrator configures the peer group with all the BGP parameters that are to be applied to many BGP peers. Actual BGP neighbors are bound to the peer group, and the network administrator applies the peer group configuration on each of the BGP sessions. BGP neighbors of a single router can be divided into several groups, each group having its own BGP parameters. Actual neighbors are then bound to the appropriate group, resulting in an optimum BGP configuration.

Example



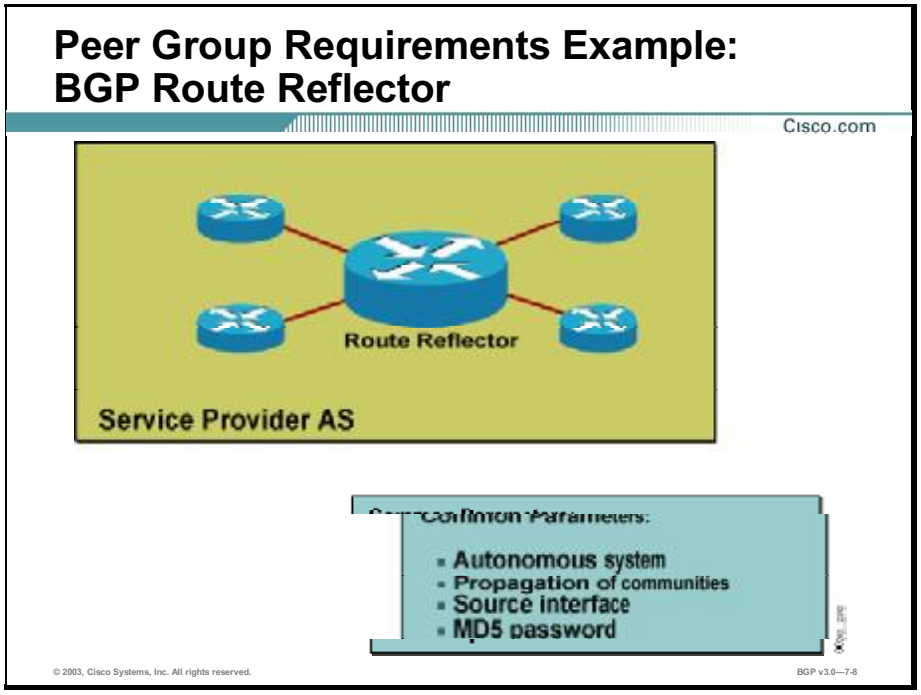
In a service provider network, a group of customer autonomous systems can be treated in the same (or a very similar) way. They are all assumed to announce local networks only. All customer autonomous systems should receive BGP updates with the same set of Internet routes, and they are all assumed to generate only a small number of prefixes.

This situation makes the neighbor configuration almost identical for each of the customers, with only a few changes specific to each neighbor.

In this scenario, the use of the peer group function is highly desirable. The network administrator can configure BGP neighbors in the customer autonomous systems using a single peer group. The administrator configures the peer group template with references to route-maps, filter-lists, prefix-lists, and the maximum number of received prefixes. Then the IP addresses of the customer routers are bound to the peer group, and the peer group configuration is applied to all of them.

 Copyright © 2003, Cisco Systems, Inc.

Example

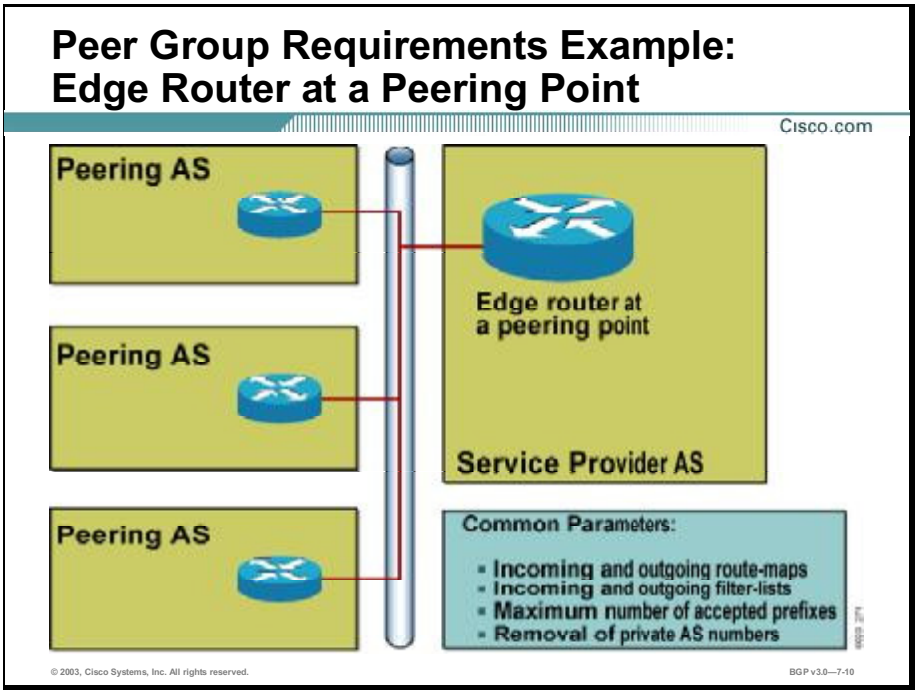


In order to apply a consistent routing policy within the entire local AS, the network administrator needs to configure each and every IBGP session identically. If a router in the AS is supplied with some information, then all the routers should be supplied with the same information. Otherwise, an inconsistent routing policy within the AS might cause inconsistent routing or application of BGP policies.

The peer group function is a good tool to make sure that all IBGP peers receive the same configuration information. The network administrator configures a peer group template with the required parameters, such as the neighbor AS number, enabling of send-community, setting of the update source to a loopback interface, and router authentication mechanisms. Then all the internal neighbor IP addresses are bound to the peer group, and the peer group configuration is applied to all of them. This approach ensures a consistent routing policy within the AS.

In a service provider network, the routers assigned as route reflectors are the routers with the largest number of IBGP sessions. These are the routers where the peer group function is most useful.

Example



The edge router located in the network where the service provider exchanges routes with other service providers is also a suitable place to use peer groups. From the edge router, the service provider AS can peer with a large number of other service providers.

All peering autonomous systems should receive the same set of routes, namely the routes local to the service provider AS and the routes received from customer autonomous systems. Also, all routes received by the service provider peering router from all peering autonomous systems are processed almost identically. The characteristic of the exchange network is the same regardless of which neighbor the routes are received from. If the peering point is an FDDI, ATM, Gigabit Ethernet, or Dynamic Packet Transport (DPT) network, the preference of using the network for packet exchange may be different. However, for each single peering point, all neighbors are reachable over the same network, and the preference is quite likely to be the same.

Additionally, a number of other parameters could be the same, such as the removal of private AS numbers before sending the routes to neighbor autonomous systems and limiting the amount of received routes. In these cases, the network administrator can apply these parameters on the peer group template before the actual IP addresses of the neighbors are bound to the peer group.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What is the purpose of using BGP peer groups?

A) BGP peer groups can be used to configure the same set of parameters for a number of BGP neighbors in a common template.

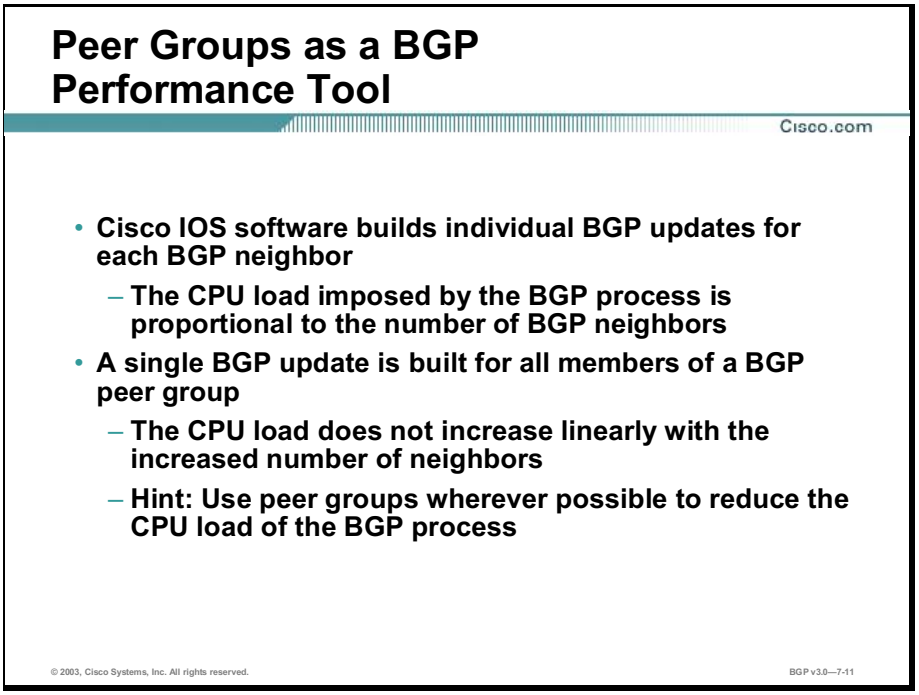
B) BGP peer groups can be used to allow anonymous BGP neighbors.

C) BGP peer groups allow EBGP peers to be configured with the same AS number and parameters.

D) BGP peer groups can be used to hide the identity of BGP peers from external neighbors.

Peer Groups as a BGP Performance Tool

This topic describes the performance benefits of using BGP peer groups.


Peer Groups as a BGP Performance Tool
Cisco.com
• Cisco IOS software builds individual BGP updates for each BGP neighbor
– The CPU load imposed by the BGP process is proportional to the number of BGP neighbors
• A single BGP update is built for all members of a BGP peer group
– The CPU load does not increase linearly with the increased number of neighbors
– Hint: Use peer groups wherever possible to reduce the CPU load of the BGP process
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-11

By default, Cisco IOS software builds BGP updates for each neighbor individually. Building BGP updates involves a number of router-CPU-–consuming tasks, including scanning the BGP table and applying a variety of outgoing filtering mechanisms (filter-lists, route-maps, and prefix-lists). These tasks imply that when a router is configured with a large number of neighbors, the CPU load grows proportionally.

However, with the use of peer groups, some of the router CPU utilization imposed by BGP update generation is significantly reduced, because the use of peer groups allows the router to run the BGP update (including all outgoing filter processing) only once for the entire peer group. The router, once finished building the BGP update, sends it to each member of the peer group. The actual TCP transmission still has to be done on a per-neighbor basis due to the connection-oriented characteristics of BGP sessions.

So, router CPU load does increase when there are more neighbors to a router due to increased TCP workload, but the use of peer groups can significantly reduce the increase. Therefore, network administrators should use peer groups whenever possible to reduce the CPU load.

Note	BGP peer groups are the fundamental BGP scalability tool and should be used in all environments where a router has a large number of BGP neighbors.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What are two performance benefits of configuring BPG peer groups on a Cisco router? (Choose two.)

A) Peer groups reduce the amount of memory required to process BGP updates between BGP peers.

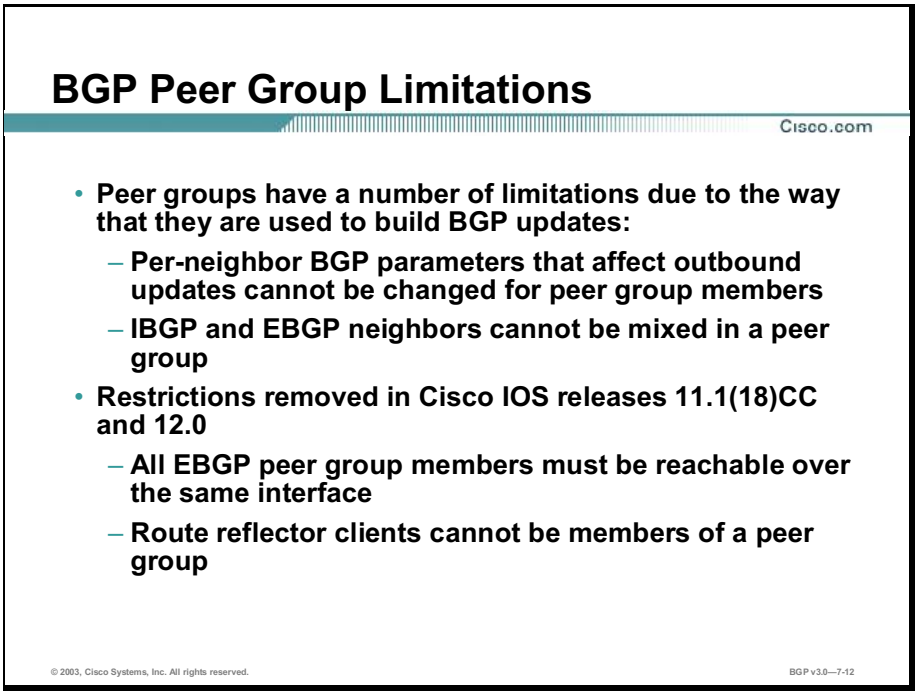
B) Peer groups reduce the amount of CPU resources required to process many BGP neighbor sessions.

C) Peer groups remove route replication requirements when routers process updates from a large number of neighbors.

D) Peer groups reduce the number of TCP sessions required to form BGP neighbor relationships.

BGP Peer Group Limitations

This topic describes the limitations of BGP peer groups.


BGP Peer Group Limitations
Cisco.com
• Peer groups have a number of limitations due to the way that they are used to build BGP updates:
– Per-neighbor BGP parameters that affect outbound updates cannot be changed for peer group members
– IBGP and EBGP neighbors cannot be mixed in a peer group
• Restrictions removed in Cisco IOS releases 11.1(18)CC and 12.0
– All EBGP peer group members must be reachable over the same interface
– Route reflector clients cannot be members of a peer group
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-12

Because the router builds only one update for all members of the same peer group, some restrictions apply to members of the peer group:

- There cannot be different outbound filters, route-maps, or other means that could possibly cause different updates to be sent to two members of the same peer group. Cisco IOS software creates only one update, which is subsequently replicated to all members. Any violation of this rule could cause unexpected results.

- External Border Gateway Protocol (EBGP) and IBGP updates are very different. EBGP updates have the AS-path attribute changed. IBGP sessions pass only the local preference attribute. The MED attribute received from a remote AS is passed onto IBGP sessions but is removed before it is sent on an EBGP session. Therefore, you cannot assign IBGP neighbors and EBGP neighbors to the same peer group because they cannot receive a replicate of the same update.

Prior to Cisco IOS release 11.1(18)CC, a route reflector client could not be a member of a peer group. When the peer group leader was a route reflector client and an update was received from it, the route reflector split-horizon rules prevented the update from being sent back to the sender. Therefore, no update was generated for any members of the peer group. When using peer groups in combination with route reflectors, make sure that all routers in the AS are running Cisco IOS releases later than 11.1(18)CC, where this restriction is lifted.

Because the router sends the same update to each of the peer group members, the next-hop BGP attribute is replicated. The receivers of the information must be able to use the same next-

 Copyright © 2003, Cisco Systems, Inc.

hop IP address, requiring the receivers to be in the same IP subnet. If two receivers are on different subnets only one of them will receive a valid next-hop attribute. The other routers will receive a next-hop IP address that is inaccessible. This restriction was also removed in Cisco IOS releases 11.1(18)CC, making BGP peer groups an ideal scalability tool.

Practice

Q1) What are two limitations of BGP peer groups on Cisco routers? (Choose two.)

A) EBGP and IBGP neighbors cannot be members of the same peer group.

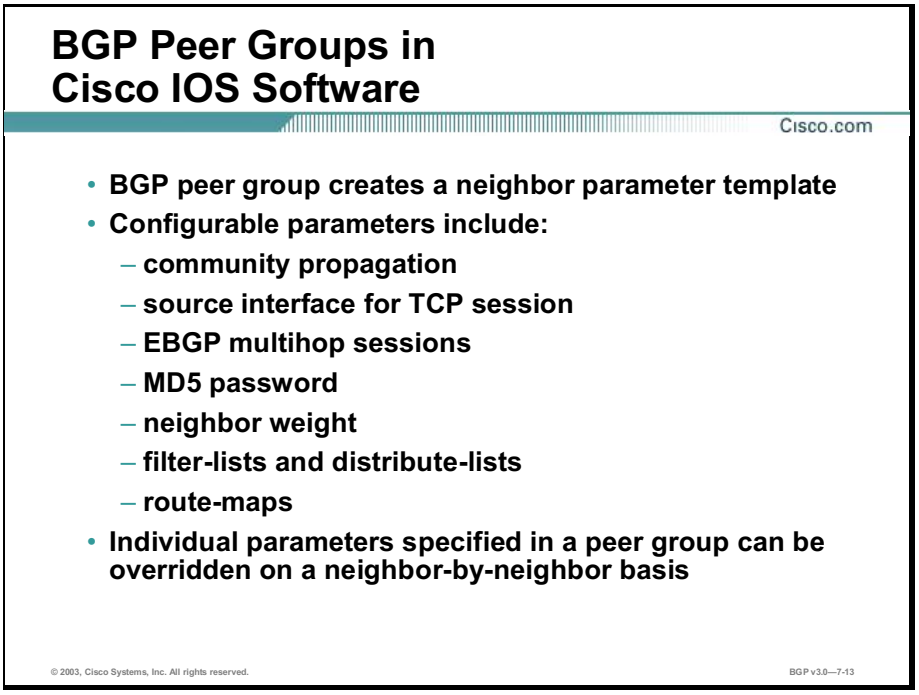
B) All routers in the peer group must belong to the same AS.

C) Peer group members cannot contain different outbound filtering mechanisms.

D) Peer group members must have the same inbound filtering mechanisms.

BGP Peer Groups in Cisco IOS Software

This topic describes the characteristics of BGP peer groups when implemented in Cisco IOS software.


BGP Peer Groups in Cisco IOS Software
Cisco.com
• BGP peer group creates a neighbor parameter template
• Configurable parameters include:
– community propagation
– source interface for TCP session
– EBGP multihop sessions
– MD5 password
– neighbor weight
– filter-lists and distribute-lists
– route-maps
• Individual parameters specified in a peer group can be overridden on a neighbor-by-neighbor basis
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-13

On a Cisco IOS router, the peer group is created as a template. The template is configured to:

- Propagate, or not propagate, the community attribute
- Use the IP address of a specific interface as the source address when opening the TCP session
- Use, or not use, the EBGP multihop function
- Use, or not use, the Message Digest 5 (MD5) authentication on the BGP sessions
- Assign a particular weight value to the routes received
- Filter out any incoming or outgoing routes based on the content of the AS-path attribute associated with the route or the network number of the route
- Pass the incoming or outgoing routes through a particular route-map

When actual neighboring routers are assigned to the peer group on a router, all of the attributes configured for the peer group are applied to all peer group members. Cisco IOS software optimizes the outgoing routes by running through the outgoing filters and route-maps only once and then replicating the results to each of the peer group members. In reality, Cisco IOS

 Copyright © 2003, Cisco Systems, Inc.

software assigns a peer group leader, for which the software generates an update, and this update is replicated to all other members of the peer group.

Some parameters configured on the peer group can be overridden by neighbor configurations, but only if the individual configurations apply on incoming updates. Outgoing updates are always prepared for the peer group leader and then replicated to the other members of the peer group.

Practice

Q1) Why must outgoing filters be applied uniformly to all members of a BGP peer group?

A) Outgoing filters are applied only once to each peer group member to conserve router CPU resources.

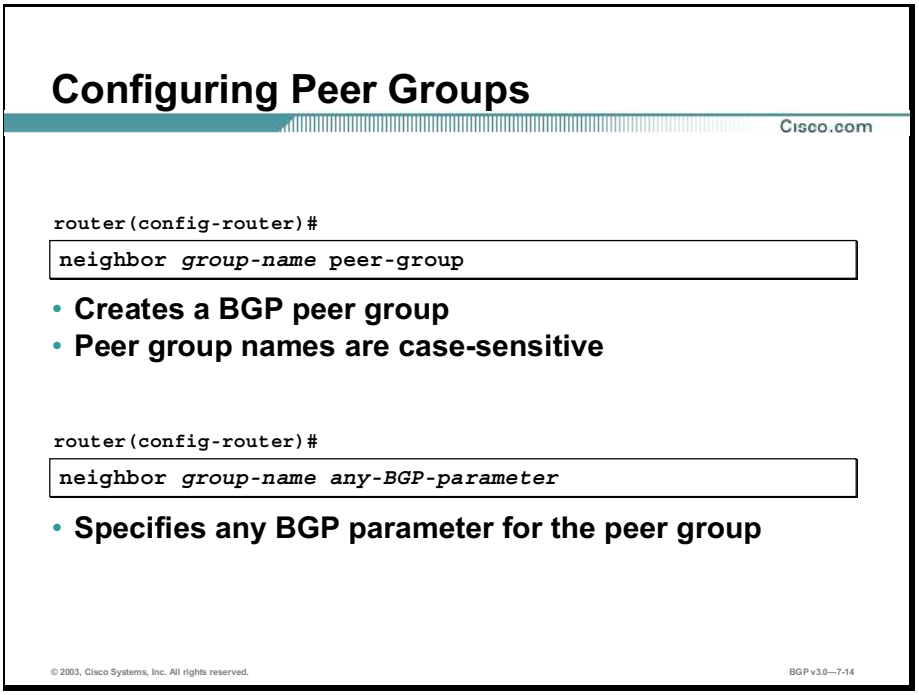
B) Outgoing filters are applied to the leader of the peer group first, and the results are replicated to other peer group members.

C) Outgoing filters are limited to a specific router-processing path that prevents them from being applied differently to different neighbors.

D) There is no restriction on applying outgoing filters to members of a peer group.

Configuring Peer Groups

This topic identifies the commands to configure BGP peer groups on Cisco IOS routers.



To configure BGP peer groups on Cisco IOS routers, perform the following steps:

- Create a BGP peer group
- Specify parameters for the BGP peer group
- Create a BGP neighbor
- Assign a neighbor into the peer group

neighbor peer-group (creating)

To create a BGP peer group, use the **neighbor peer-group** router configuration command.

neighbor *peer-group-name* **peer-group**

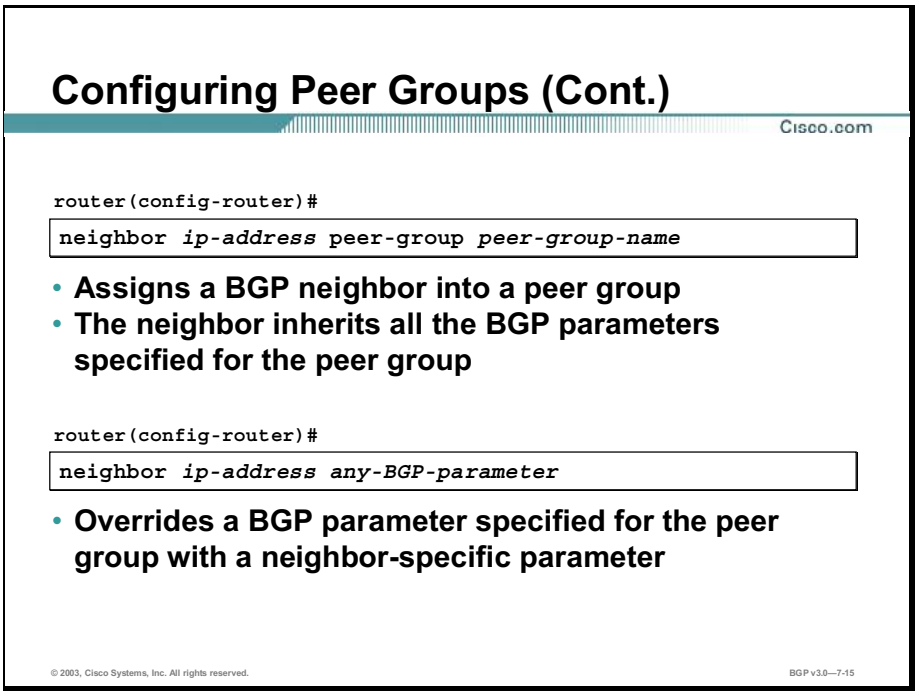
To remove the peer group and all of its members, use the **no** form of this command.

no neighbor *peer-group-name* **peer-group**

After you have created a peer group using the **neighbor peer-group** command, you can configure it with the **neighbor** commands. By default, members of the peer group inherit all the configuration options of the peer group. You can also configure members to override the options that do not affect outbound updates.

 Copyright © 2003, Cisco Systems, Inc.

Peer group members will always inherit the following configuration options: **remote-as** (if configured), **version**, **update-source**, **out-route-map**, **out-filter-list**, **out-dist-list**, **minimum-advertisement-interval**, and **next-hop-self**. All peer group members will inherit changes made to the peer group.


Configuring Peer Groups (Cont.)
Cisco.com
router(config-router)#
neighbor ip-address peer-group peer-group-name
• Assigns a BGP neighbor into a peer group
• The neighbor inherits all the BGP parameters specified for the peer group
router(config-router)#
neighbor ip-address any-BGP-parameter
• Overrides a BGP parameter specified for the peer group with a neighbor-specific parameter
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-15

neighbor peer-group (assigning members)

To configure a BGP neighbor to be a member of a peer group, use the **neighbor peer-group** router configuration command.

neighbor *ip-address* **peer-group** *peer-group-name*

To remove the neighbor from the peer group, use the **no** form of this command.

no neighbor *ip-address* **peer-group** *peer-group-name*

Syntax Description

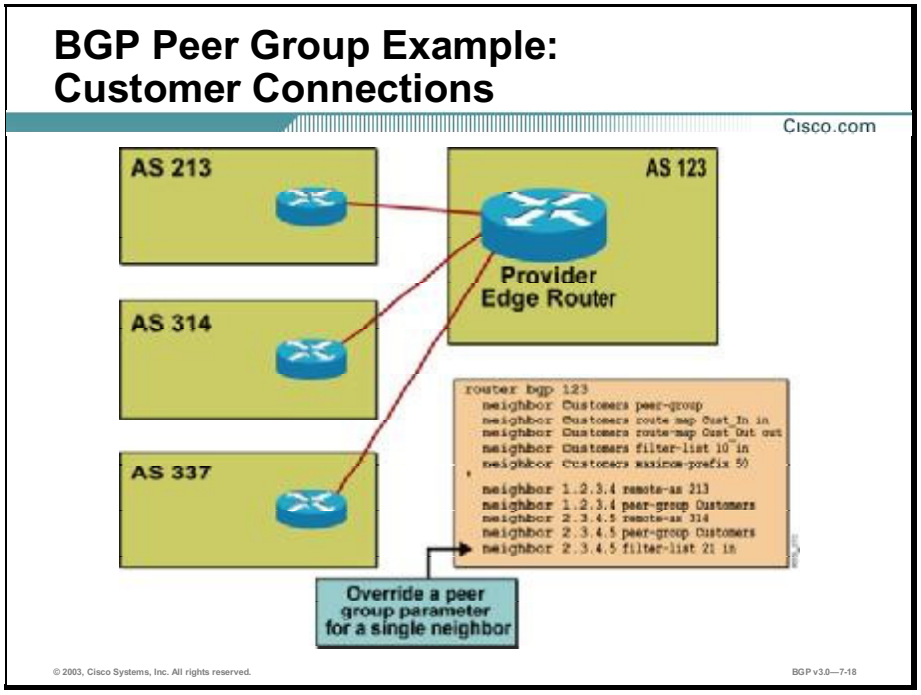
ip-address	IP address of the BGP neighbor who belongs to the peer group specified by the tag
peer-group-name	Name of the BGP peer group to which this neighbor belongs

After you have assigned an actual neighbor to be a member of the peer-group, all configurations made to the peer-group template are then applied to all the neighbors assigned to that peer group.

Through configuration, peer group configurations may be overridden on an individual neighbor basis provided the changes apply only to incoming updates. Remember that outgoing updates are prepared only once and then replicated.

 Copyright © 2003, Cisco Systems, Inc.

Example



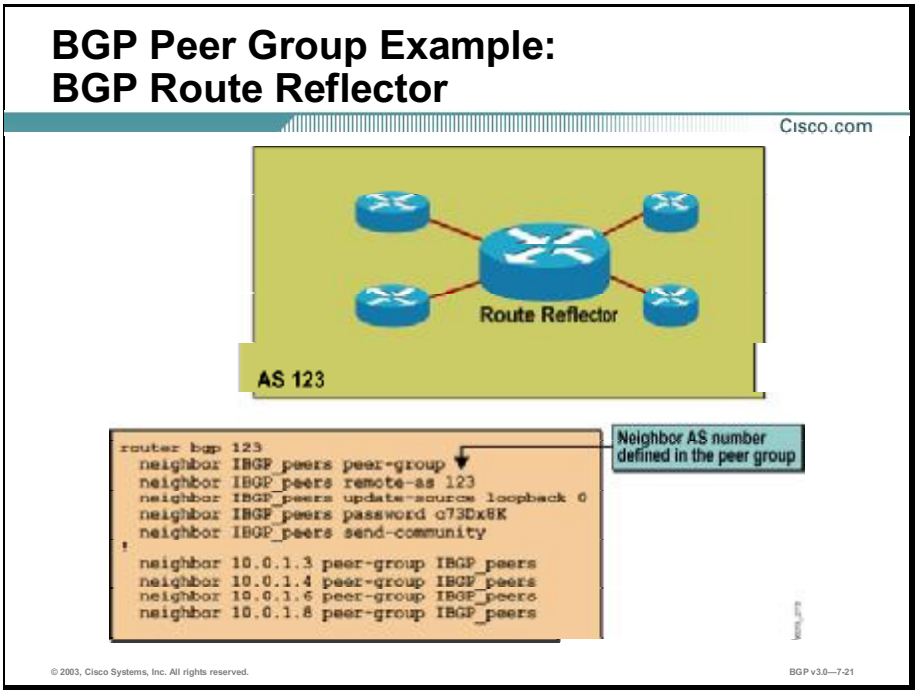
In this example, the router in AS 123 is being configured with a peer group named "Customers." This peer group is used for all customers of the service provider because they share an almost identical routing policy. The peer group is first created as a template, which is configured with an incoming route-map named *Cust_In* and an outgoing route-map named *Cust_Out*, as well as an incoming AS-path filter-list number of 10. The peer group is also configured with a maximum limit of 50 received prefixes.

Then neighbors in AS 213 and AS 314 are assigned to the peer group. These additions mean that the router in AS 123 will attempt to open BGP sessions with those routers. If the BGP sessions succeed, the route-maps *Cust_In* and *Cust_Out* will be used with both neighbors on incoming and outgoing routes, respectively. The maximum number of received prefixes configured in the "Customer" peer group will also be applied to both neighbors.

Filter-list 10 will be used to filter out any incoming routes from peer group members unless otherwise specified. However, in the case of the neighbor in AS 314, the individual configuration of filter-list 21 will override the peer group configuration, and the AS-path access-list number 21 will be used instead.

The peer group is a very powerful tool when network administrators are dealing with a large number of neighbors with almost identical configurations. However, if any of the customers require routing information that differs from what other members of the "Customer" peer group receive, then that neighbor must be removed from the peer group and configured as an individual neighbor.

Example



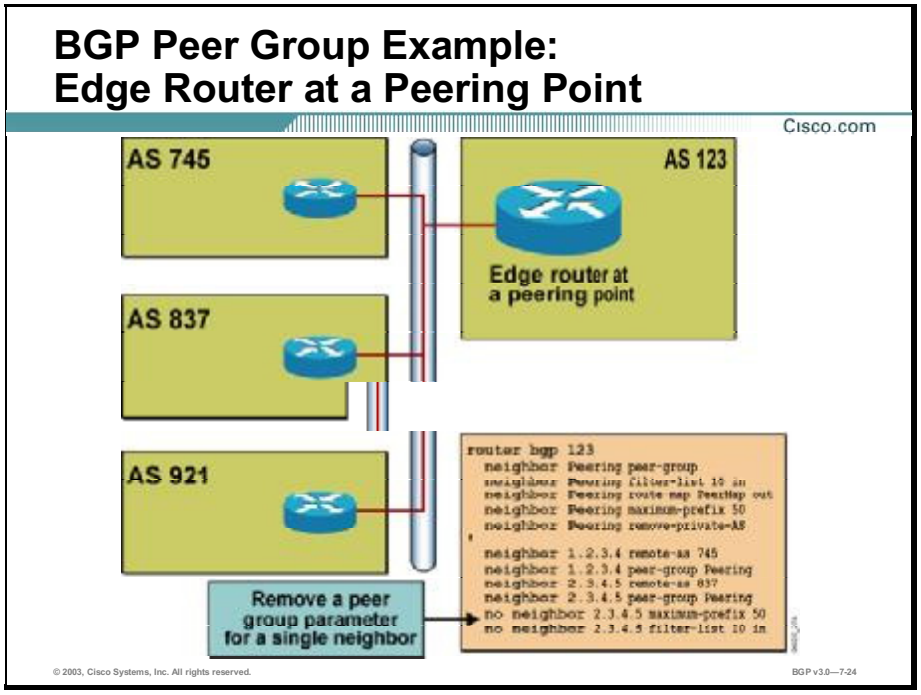
In a large AS, some routers may have a large number of IBGP sessions. In this example, a router acting as a BGP route reflector has four IBGP neighbors. The peer group named *IBGP_peers* is created to handle all of the IBGP sessions. The peer group is created and configured with the remote-AS, update-source, MD5 authentication, and community passing parameters. When the actual neighbors are configured as members of the peer group, all these configuration parameters will apply to all of the neighbors.

In the IBGP case, the remote AS can also be configured as a part of the peer group configuration because the AS number is the same for each of the peer group members.

The peer group is a very powerful tool when you are dealing with a large number of IBGP neighbors because all of them should have the same configuration to ensure a consistent AS-wide routing policy.

 Copyright © 2003, Cisco Systems, Inc.

Example



In this example, the router in AS 123 is being configured with a peer group named *Peering*. This peer group is used for all peer providers because they share an almost identical routing policy. The peer group is first created as a template, which is configured with an incoming AS-path filter-list (list 10) and an outgoing route-map named *PeerMap*. The maximum number of received prefixes is also set in the peer group to 50. The peer group has also been configured to remove private AS numbers (AS numbers in the range 64512 to 65535 inclusive) from all AS paths before the routes are sent to the peer AS.

The neighbors in AS 745 and in AS 837 are then assigned to the peer group, meaning that the router in AS 123 will attempt to open BGP sessions with those routers. If the BGP sessions are successfully established, the route-map filter-list 10 and the route-map *PeerMap*, as configured in the peer group will be applied to incoming and outgoing routes from both neighbors, respectively.

As defined in the router configuration, filter-list 10 filters out any incoming routes from peer group members unless otherwise specified. However, in the case of the neighbor in AS 837, the individual configuration of **no filter-list 10** will override the peer group configuration, and thus, the filter-list will not be used for this neighbor. The limitation on the number of received routes from AS 837 is also removed from the neighbor in AS 837.

The peer group is a very powerful tool when dealing with a large number of neighbors with almost identical configurations. However, if any of the customers assigned to the peer group require routing information that is different from other members of the peer group, then that neighbor must be removed from the peer group and configured individually.

Practice

Q1) What are three steps required to properly configure BGP peer groups on Cisco routers? (Choose three.)

A) Specify parameters for the BGP peer group.

B) Create a BGP peer group.

C) Enable the peer group by clearing the BGP session.

D) Assign a neighbor into the peer group.

Q2) What happens if a neighboring router that is configured as a member of a peer group does not support BGP peer groups?

A) The BGP peer group negotiation fails, and the neighbor is removed from the peer group.

B) The BGP session negotiation fails; it oscillates between Idle and Active states.

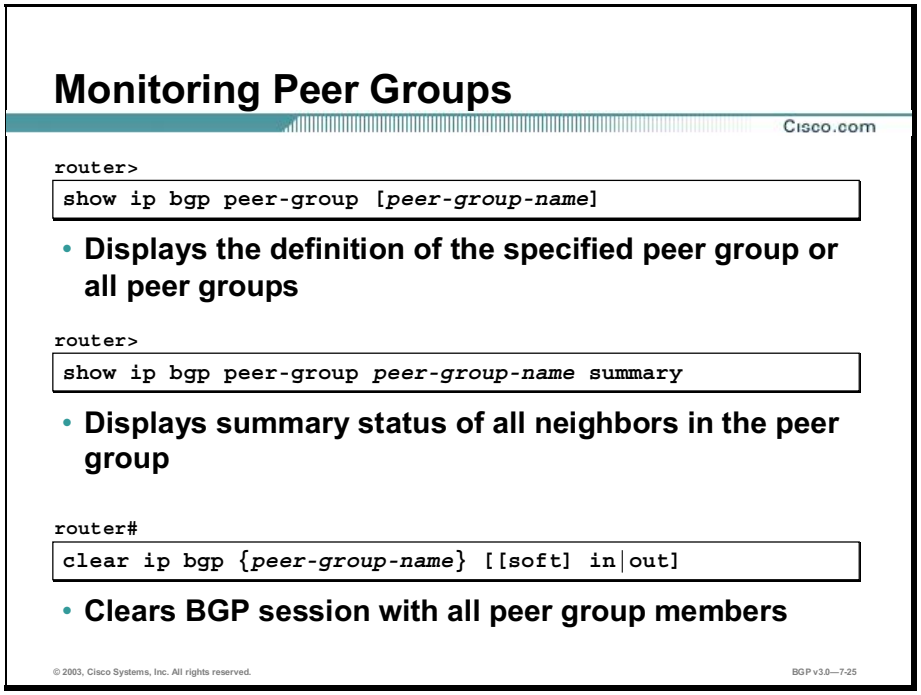
C) The local router will reject the peer group configuration.

D) The BGP peer group functionality does not require any support by the neighbor.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring Peer Groups

This topic lists the Cisco IOS commands required to monitor BGP peer groups.


Monitoring Peer Groups
Cisco.com
router>
show ip bgp peer-group [peer-group-name]
• Displays the definition of the specified peer group or all peer groups
router>
show ip bgp peer-group peer-group-name summary
• Displays summary status of all neighbors in the peer group
router#
clear ip bgp {peer-group-name} [[soft] in|out]
• Clears BGP session with all peer group members
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-25

show ip bgp peer-group

To display information about BGP peer groups, use the **show ip bgp peer-group** EXEC command.

show ip bgp peer-group [*peer-group-name*] [**summary**]

Syntax Description

peer-group-name	(Optional) Displays information about that specific peer group
summary	(Optional) Displays a summary of the status of all the members of a peer group

clear ip bgp

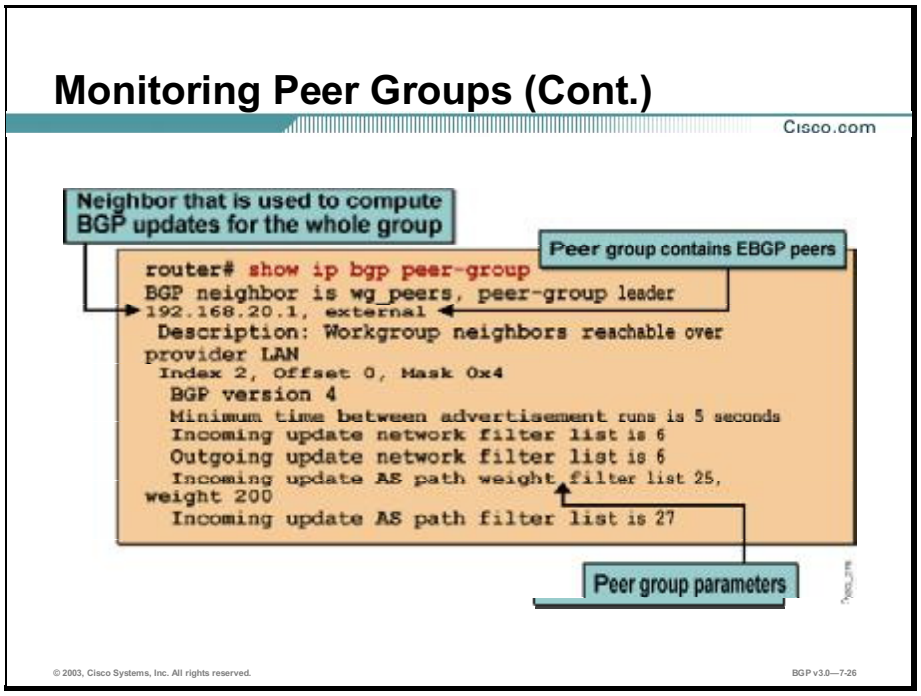
To reset the BGP sessions with all the members of a peer group, use the **clear ip bgp** EXEC command.

clear ip bgp {***** | *neighbor-address* | *peer-group-name*} [**soft** [**in** | **out**]]

Syntax Description

*	Resets all current BGP sessions.
neighbor-address	Resets only the identified BGP neighbor.
peer-group-name	Resets the specified BGP peer group.
soft	(Optional) Soft reconfiguration.
in | **out**	(Optional) Triggers inbound or outbound soft reconfiguration. If you do not specify the **in** or **out** option, both inbound and outbound soft reconfiguration are triggered.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Peer Groups (Cont.)
Cisco.com
Neighbor that is used to compute BGP updates for the whole group
Peer group contains EBGP peers
router# show ip bgp peer-group
BGP neighbor is wg_peers, peer-group leader
192.168.20.1, external
Description: Workgroup neighbors reachable over provider LAN
Index 2, Offset 0, Mask 0x4
BGP version 4
Minimum time between advertisement runs is 5 seconds
Incoming update network filter list is 6
Outgoing update network filter list is 6
Incoming update AS path weight filter list 25, weight 200
Incoming update AS path filter list is 27
Peer group parameters
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-26

In this example, the **show ip bgp peer-group** command displays information about the peer group named *wg_peers*. One of the peer group members (192.168.20.1) has been selected as the peer group leader, meaning that all outgoing BGP updates are processed as if they were being sent to this neighbor. Thereafter, the update is replicated on all the BGP sessions to the other members in the *wg_peers* peer group.

A peer group should have only IBGP members or EBGP members. In the example above, the members are EBGP neighbors.

All parameters configured for the peer group are listed in the **show ip bgp peer-group** command output. In the example, the peer group *wg_peers* has been configured as:

```
neighbor wg_peers description Workgroup neighbors reachable over
provider LAN
neighbor wg_peers distribute-list 6 in
neighbor wg_peers distribute-list 6 out
neighbor wg_peers filter-list 25 weight 200
neighbor wg_peers filter-list 27 in
```

Monitoring Peer Groups (Cont.)

Cisco.com

```
router# show ip bgp peer-group wg_peers summary
BGP table version is 56, main routing table version 56
51 network entries (51/153 paths) using 10560 bytes of memory
18 BGP path attribute entries using 2296 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State
192.168.20.1    4     1       0       0        0    0    0 never    Active
192.168.20.2    4     2       0       0        0    0    0 never    Active
192.168.20.3    4     3       0       0        0    0    0 never    Active
192.168.20.5    4     5      53      61       56    0    0 00:00:52
192.168.20.20   4    20      62      59       56    0    0 00:00:44
192.168.20.22   4    22      54      54       56    0    0 00:00:44
```

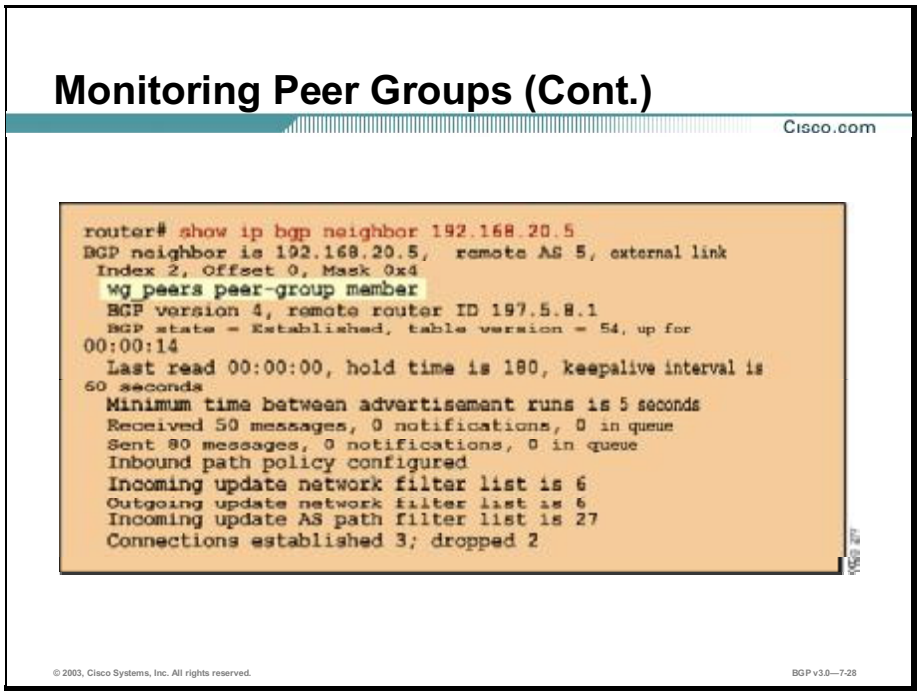
- **The printout is identical to a show ip bgp summary printout but displays only neighbors that are members of the specified peer group**

© 2003, Cisco Systems, Inc. All rights reserved.

The **show ip bgp peer-group** *peer-group-name* **summary** command is used in this example to display only the summary status information about the neighbors who are members of the peer group *wg_peers*.

Note	The **show ip bgp summary** command is described in the BGP Overview module.

 Copyright © 2003, Cisco Systems, Inc.


Monitoring Peer Groups (Cont.)
Cisco.com
router# show ip bgp neighbor 192.168.20.5
BGP neighbor is 192.168.20.5, remote AS 5, external link
Index 2, Offset 0, Mask 0x4
wg_peers peer-group member
BGP version 4, remote router ID 197.5.8.1
BGP state = Established, table version = 54, up for
00:00:14
Last read 00:00:00, hold time is 180, keepalive interval is
60 seconds
Minimum time between advertisemant runs is 5 seconds
Received 50 messages, 0 notifications, 0 in queue
Sent 80 messages, 0 notifications, 0 in queue
Inbound path policy configured
Incoming update network filter list is 6
Outgoing update network filter list is 6
Incoming update AS path filter list is 27
Connections established 3; dropped 2
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-28

The **show ip bgp neighbor** command displays additional information about BGP neighbors that are members of a peer group.

In the example above, the membership in the peer group *wg_peers* is indicated, in addition to the other information as displayed by the command.

Practice

Q1) What command do you use to display the summary status of all neighbors in a peer group?

A) **show ip bgp**

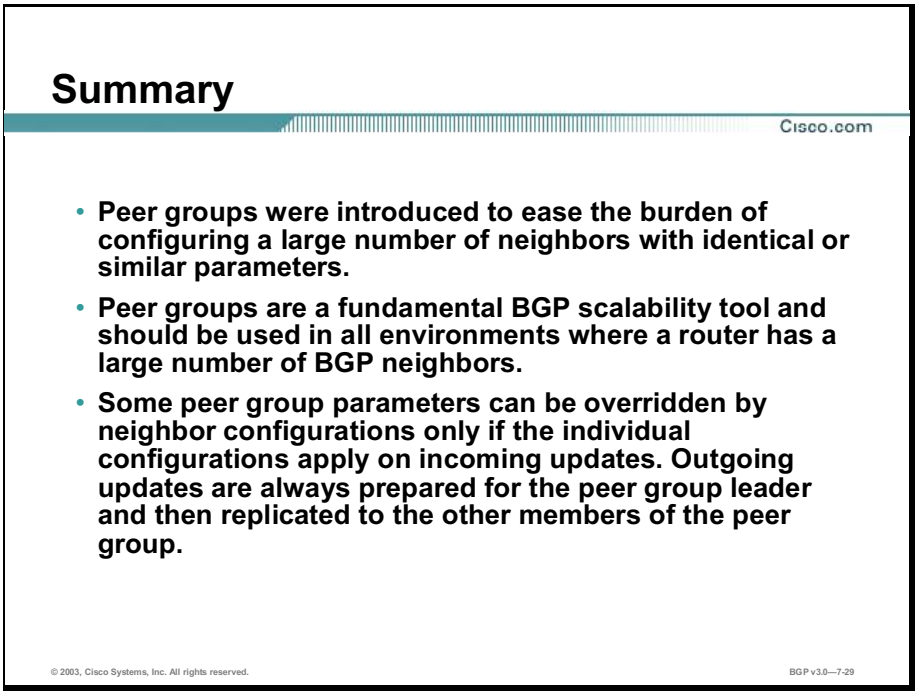
B) **show peer-group summary**

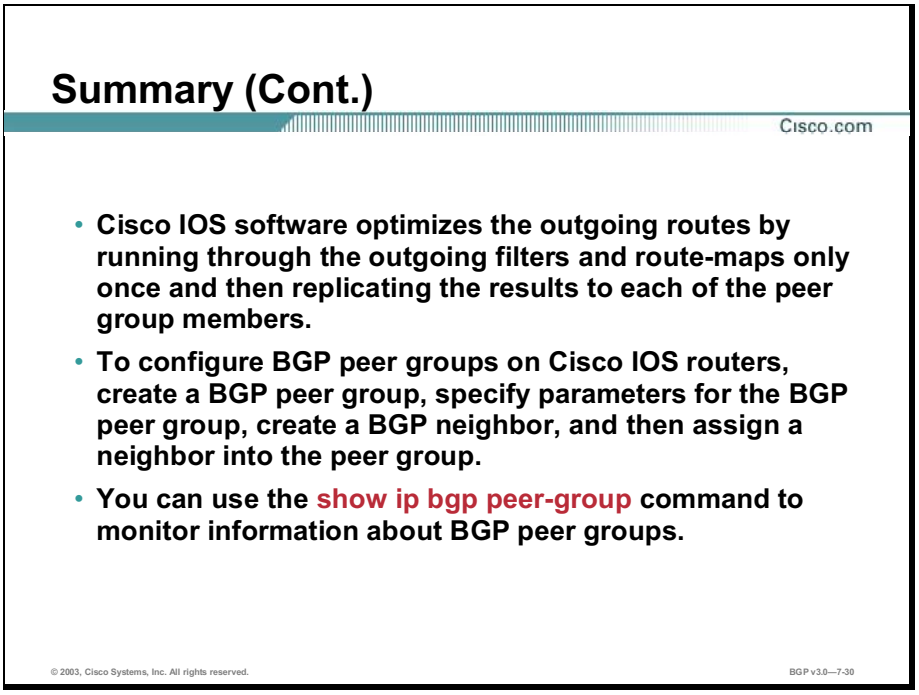
C) **show ip bgp neighbor**

D) **show ip bgp peer-group summary**

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Peer groups were introduced to ease the burden of configuring a large number of neighbors with identical or similar parameters.
• Peer groups are a fundamental BGP scalability tool and should be used in all environments where a router has a large number of BGP neighbors.
• Some peer group parameters can be overridden by neighbor configurations only if the individual configurations apply on incoming updates. Outgoing updates are always prepared for the peer group leader and then replicated to the other members of the peer group.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-29


Summary (Cont.)
Cisco.com
• Cisco IOS software optimizes the outgoing routes by running through the outgoing filters and route-maps only once and then replicating the results to each of the peer group members.
• To configure BGP peer groups on Cisco IOS routers, create a BGP peer group, specify parameters for the BGP peer group, create a BGP neighbor, and then assign a neighbor into the peer group.
• You can use the show ip bgp peer-group command to monitor information about BGP peer groups.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-30

 Copyright © 2003, Cisco Systems, Inc.

Next Steps

After completing this lesson, go to:

- BGP Route Dampening lesson

References

For additional information, refer to these resources:

- For more information on BGP peer groups, refer to “BGP Peer Groups” at the following URL: http://www.cisco.com/warp/public/459/29.html

- For more information on configuring BGP peer groups, refer to “Configuring BGP” at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/np1_c/1cprt1/1cbgp.htm#xtocid45

Laboratory Exercise: BGP Peer Groups

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will use BGP peer groups to share common configuration parameters between multiple BGP peers.

After completing this exercise, you will be able to:

- Configure BGP peer groups
- Monitor the operation of BGP peer groups

 Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor *peer-group-name* **peer-group**	To create a BGP or multiprotocol BGP peer group.
neighbor *ip-address* **peer-group** *peer-group-name*	To configure a BGP neighbor to be a member of a peer group.
neighbor *peer-group-name* **maximum-prefix** *maximum* [**threshold**] [**warning-only**][**restart** *restart-interval*]	Specify the allowed maximum number of prefixes.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp neighbors *neighbor*	Show detailed information about a specific neighbor.
show ip bgp peer-group [*peer-group-name*] [**summary**]	To display information about BGP peer groups.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- After working with the “Cheap” provider, you have finally resolved the maximum-prefix violation problem that your network was experiencing. You have decided to increase the number of prefixes allowed and to make prefix limiting a universally applied policy for all peers.
- In this exercise, you will configure router WGxR1 to terminate the neighbor session between you and your upstream providers if you receive more than 50 BGP prefixes from them.
- You will use BGP peer groups to implement the policy to ease administration when more peers are added and to minimize router CPU utilization.
- Figure 1 shows the connectivity that is established between your AS and the two service providers “Good” and “Cheap.”

Copyright © 2003, Cisco Systems, Inc.

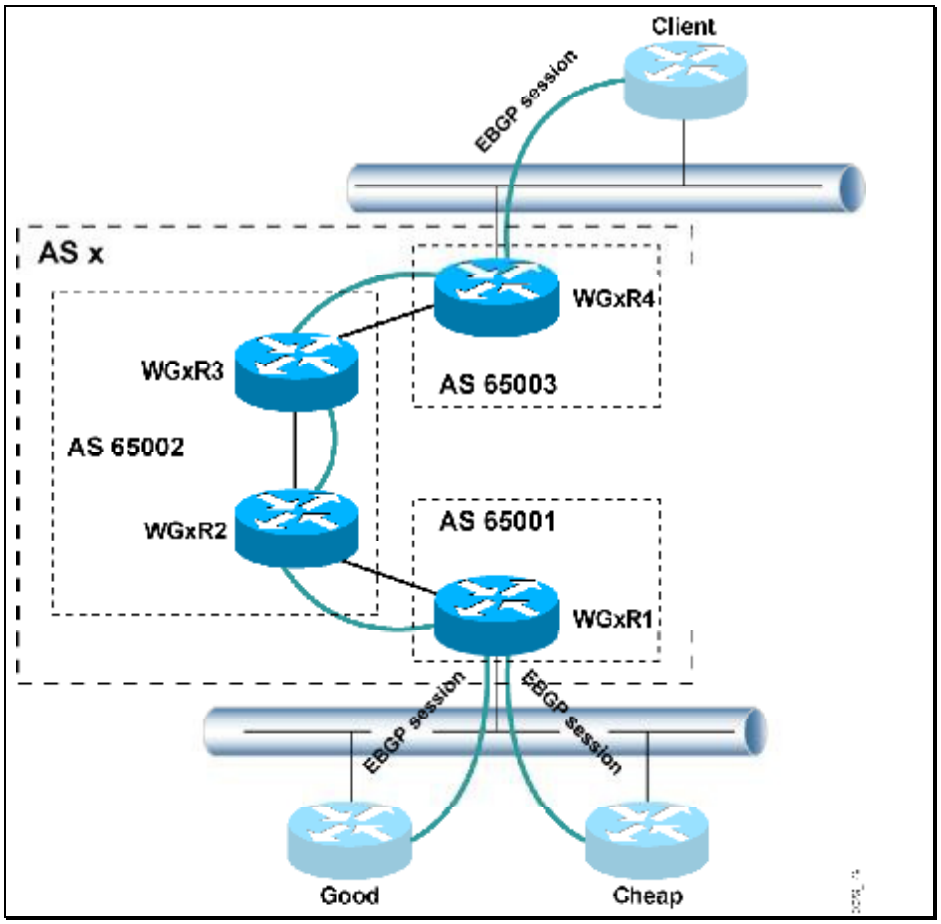


Figure 1: Connections to two different service providers

Exercise Procedure

Complete these steps:

Create a BGP peer group

Step 1 Create a BGP peer group on your WGxR1 router to set a limit to the number of prefixes that your router accepts. BGP sessions should be terminated if the number of received prefixes exceeds 50.

Assign a neighbor into a peer group

Step 2 Configure the neighbor connections for the service provider routers “Good” and “Cheap” to be members of the peer group.

 Copyright © 2003, Cisco Systems, Inc.

Exercise Verification

You have completed this exercise when you attain these results:

- On WGxR1, use the **show ip bgp peer-group** *peer-group-name* command with the **summary** keyword to verify that you have configured the correct neighbors as members of the peer group.

```
WG1R1#sh ip bgp peer-group SvcPro sum
BGP router identifier 197.1.8.1, local AS number 65001
BGP table version is 355, main routing table version 355
43 network entries and 66 paths using 6547 bytes of memory
20 BGP path attribute entries using 1040 bytes of memory
19 BGP AS-PATH entries using 456 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP activity 84/1135 prefixes, 319/253 paths, scan interval 15 secs
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.20.20   4    20   13979    8418      355    0    0 00:15:29
23
192.168.20.22   4    22     412     397      355    0    0 00:07:38
23
```

- On WGxR1, use the **show ip bgp peer-group** command to verify that the parameters of the peer group have been properly configured.

```
WG1R1#sh ip bgp peer-group
BGP peer-group is SvcPro
  BGP version 4
  Default minimum time between advertisement runs is 30 seconds
 For address family: IPv4 Unicast
  BGP neighbor is SvcPro, peer-group external, members:
     192.168.20.20   192.168.20.22
  Index 1, Offset 0, Mask 0x2
  Inbound soft reconfiguration allowed
  Update messages formatted 38, replicated 13
```

- Inspect the prefixes that your router is receiving from both “Good” and “Cheap” with the **show ip bgp neighbors** command. Be sure to use the **received-routes** keyword.

```
wg1r1#sh ip bgp neighbors 192.168.20.20 received-routes
BGP table version is 355, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.0.0.0         192.168.20.233           0             0 20 i
```

```
*> 128.20.0.0       192.168.20.20           0             0 20 i
*> 128.20.12.0/24   192.168.20.20           0             0 20 i
*> 128.22.0.0       192.168.20.20           0             0 20 i
*  128.22.12.0/24   192.168.20.22                         0 20 22 i
*  128.26.0.0       192.168.20.20           0             0 20 42 26
i
*> 128.37.0.0       192.168.20.20           0             0 20 42 37
i
*> 128.42.0.0       192.168.20.20           0             0 20 42 i
*  128.51.0.0       192.168.20.20           0             0 20 42 26
51 i
*> 128.213.0.0      192.168.20.20           0             0 20 213 i
*  128.214.0.0      192.168.20.22                         0 20 22
214 i
*> 192.20.11.0      192.168.20.20           0             0 20 i
*> 192.20.12.0/30   192.168.20.20           0             0 20 i
*> 192.22.11.0      192.168.20.20           0             0 20 i
*  192.22.12.0/30   192.168.20.22                         0 20 22 i
*  192.26.11.0      192.168.20.20           0             0 20 42 26
i
*> 192.37.11.0      192.168.20.20           0             0 20 42 37
i
*> 192.42.11.0      192.168.20.20           0             0 20 42 i
*  192.51.11.0      192.168.20.20           0             0 20 42 26
51 i
*> 192.213.11.0     192.168.20.20           0             0 20 213 i
*  192.214.11.0     192.168.20.22                         0 20 22
214 i
*> 200.20.0.0/16    192.168.20.20           0             0 20 i
*  200.22.0.0/16    192.168.20.22                         0 20 22 i
Total number of prefixes 23
```

- Clear the neighbor session between your WGxR1 and router “Cheap” afterward and ensure that the neighbor session remains active.

Answer this question:

Q1) If the number of prefixes received from both providers exceeds 50, what single command can you use to reset the BGP sessions on both routers?

 Copyright © 2003, Cisco Systems, Inc.

BGP Route Dampening

Overview

The Border Gateway Protocol (BGP) was designed to scale to meet the interdomain routing goals of large networks. To continue to scale to meet the growing routing requirements of the Internet, it is important that BGP processing overhead be reduced. Two factors contributing most to BGP processing requirements are the addition or deletion of a prefix entry (in the BGP routing table) and the BGP route selection process. One way of reducing BGP processing requirements is to stabilize the state of the BGP routing table by reducing the number of routing table state changes propagated by BGP.

Route dampening is a BGP feature designed to reduce BGP processing requirements by minimizing the propagation of unstable routes to BGP peers. Autonomous system (AS) border routers, in any BGP implementation, cannot rely upon external peers to sufficiently shield the AS from route table instability. Route dampening allows route instability to be contained at an AS border router that borders the instability.

This lesson describes the purpose and operation of the route-dampening feature on Cisco IOS® routers. Also discussed in this lesson are the Cisco IOS commands required to enable route dampening, modify default dampening parameters, and to release a route that has been suppressed due to dampening. The Cisco IOS commands used to monitor route dampening are also discussed.

Importance

Even when a BGP implementation is correctly configured and highly robust, the performance of the routing process on any given router is limited. Limiting the propagation of unstable routes, specifically when not beneficial to the network, becomes an important issue because it reduces the processing requirements of the router that is forced to process route table state changes. Route dampening is an important scalability feature for BGP routers.

Objectives

Upon completing this lesson, you will be able to:

- Describe the purpose of BGP route dampening
- Describe the operation of BGP route dampening
- Identify the Cisco IOS commands required to configure BGP route dampening
- Identify the Cisco IOS commands required to release dampened routes
- Identify the Cisco IOS commands required to monitor BGP route-dampening operation

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

Outline

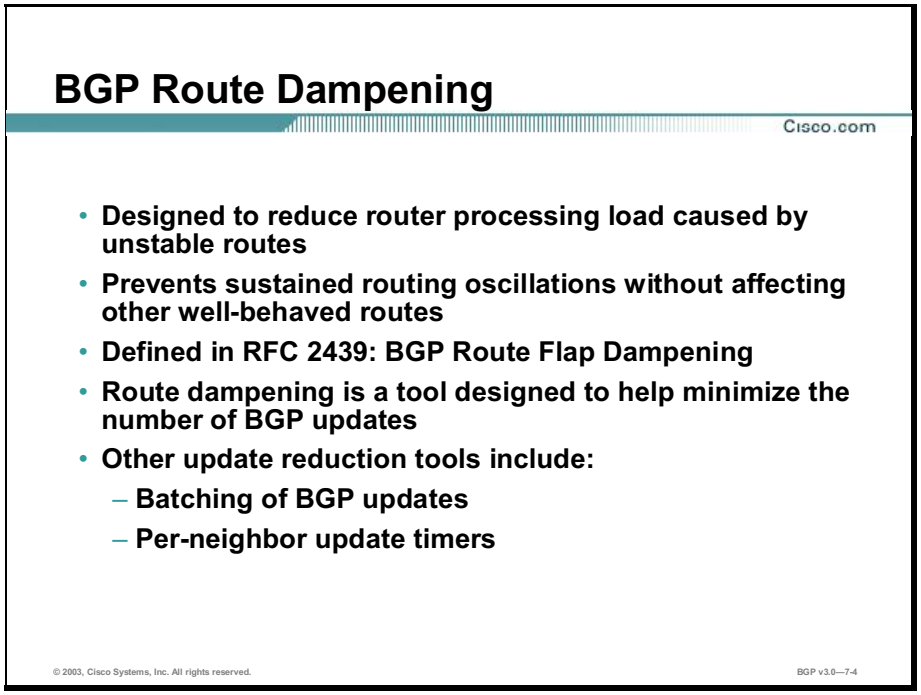
This lesson includes these topics:

- Overview
- BGP Route Dampening
- Route-Dampening Operation
- Configuring BGP Route Dampening
- Releasing Dampened Routes
- Monitoring Route Dampening
- Summary
- Assessment (Lab): BGP Route Dampening

 Copyright © 2003, Cisco Systems, Inc.

BGP Route Dampening

This topic describes the purpose of BGP route dampening.

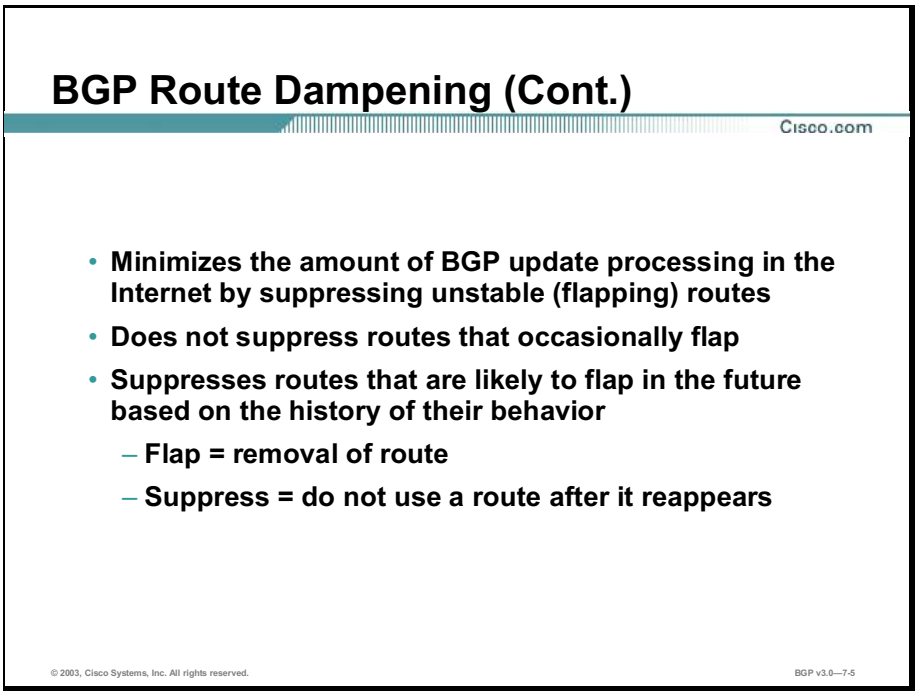

BGP Route Dampening
Cisco.com
• Designed to reduce router processing load caused by unstable routes
• Prevents sustained routing oscillations without affecting other well-behaved routes
• Defined in RFC 2439: BGP Route Flap Dampening
• Route dampening is a tool designed to help minimize the number of BGP updates
• Other update reduction tools include:
– Batching of BGP updates
– Per-neighbor update timers
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-4

BGP is the only routing protocol designed for large internetworks with the specific intention of carrying a large number of prefixes. There are several mechanisms built into BGP that ensure maximum router stability.

For example, a BGP router does not forward BGP routing updates immediately after receiving them. Every time a BGP router sends an update, it starts a 5-second timer for internal neighbors and a 30-second timer for external neighbors. No new updates can be sent until that timer expires. The result is that if a router contained a link that was flapping repeatedly (available, then unavailable, then available, then unavailable, and so on) at a rate of once per second, external routers see the flap at a much slower rate. Routers external to the source of the flap are not forced to recalculate the best path every second but, at most, every 30 seconds.

Reducing the rate at which neighboring routers process flapping routes assists in reducing the requirements to process the BGP update. However, routers that process routing updates for unstable routes are still wasting resources determining the best route to the destination. Because the unstable route is oscillating between up and down, each route update that a router receives causes it to process the unstable route all over again. A better approach is to remove the update about the route until it can be guaranteed that the destination is more stable.

With this goal in mind, an additional BGP scalability mechanism called route flap dampening was created to reduce route update processing requirements by suppressing unstable routes.


BGP Route Dampening (Cont.)
Cisco.com
• Minimizes the amount of BGP update processing in the Internet by suppressing unstable (flapping) routes
• Does not suppress routes that occasionally flap
• Suppresses routes that are likely to flap in the future based on the history of their behavior
– Flap = removal of route
– Suppress = do not use a route after it reappears
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-5

Most service providers hold routing information for the entire Internet. Therefore, a flapping link somewhere in the Internet can cause all routers in the Internet to keep processing changes because of one single link. If, however, one of the autonomous systems in the Internet implements route dampening, the flapping network is suppressed. The route is not propagated further to other autonomous systems until the configured rules of route dampening allow it.

A flap is defined as a route that is repeatedly available, then unavailable, then available, then unavailable, and so on. If a route flaps once or twice, it is typically not considered a flap from an administrative perspective. If the flapping happens more often, however, there is probably something wrong with the destination and it should be suppressed. The BGP router stores a suppressed route in the BGP table but does not consider it in the BGP path selection process and does not therefore propagate it to other BGP neighbors or use it for data forwarding.

Practice

Q1) What two mechanisms are built into BGP to make it more scalable by reducing the route-processing requirements of BGP routers? (Choose two.)

A) split horizon

B) route dampening

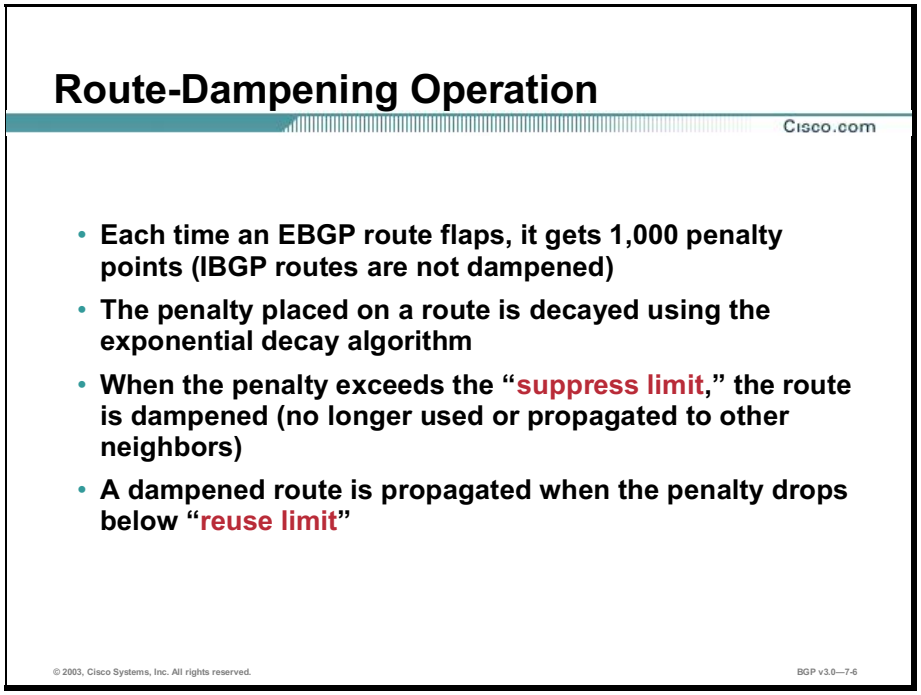
C) synchronization

D) per-neighbor update timers

 Copyright © 2003, Cisco Systems, Inc.

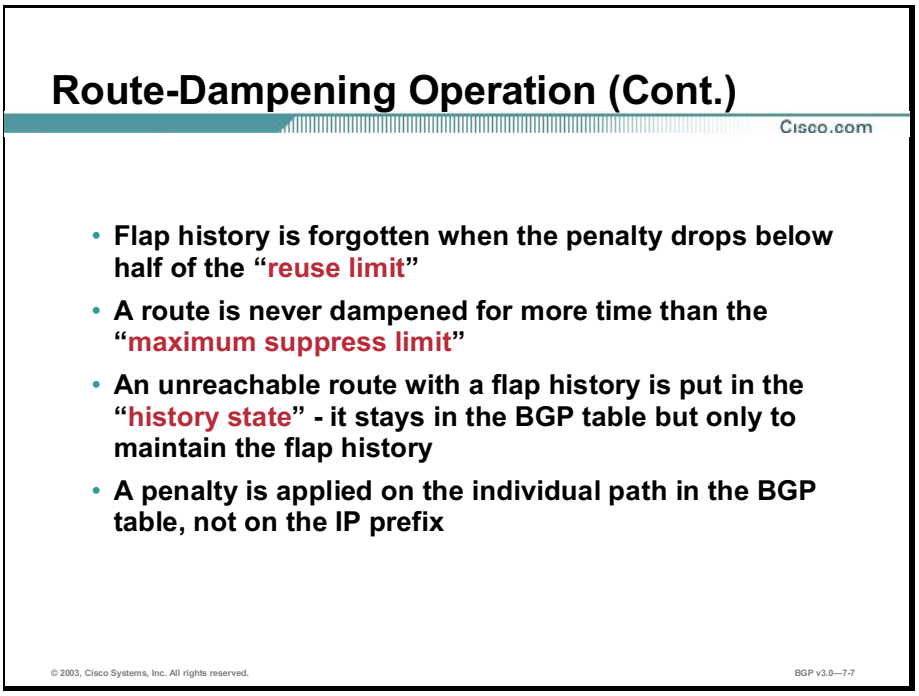
Route-Dampening Operation

This topic describes the operation of BGP route dampening.


Route-Dampening Operation
Cisco.com
• Each time an EBGP route flaps, it gets 1,000 penalty points (IBGP routes are not dampened)
• The penalty placed on a route is decayed using the exponential decay algorithm
• When the penalty exceeds the "suppress limit," the route is dampened (no longer used or propagated to other neighbors)
• A dampened route is propagated when the penalty drops below "reuse limit"
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-6

A BGP router with route dampening enabled keeps track of all routes (even those that are unreachable) so that it can recall the penalties assigned to each route. Every time a route flap occurs, the flapping route receives 1,000 penalty points. The penalty is gradually decreased using a decaying algorithm. If a route flaps several times, it will be penalized (gain enough penalty points) and subsequently reach and exceed the suppress limit.

Any route reaching the suppress limit is no longer forwarded to other neighbors until the assigned penalty is once again below the reuse limit. An exponential decay algorithm reduces penalty points applied to a flapping route. After the number of penalty points assigned to a route falls below the reuse limit, the BGP router once again advertises the route.


Route-Dampening Operation (Cont.)
Cisco.com
• Flap history is forgotten when the penalty drops below half of the "reuse limit"
• A route is never dampened for more time than the "maximum suppress limit"
• An unreachable route with a flap history is put in the "history state" - it stays in the BGP table but only to maintain the flap history
• A penalty is applied on the individual path in the BGP table, not on the IP prefix
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-7

A router stops tracking penalty points once they are below half of the reuse limit.

The maximum suppress limit defines the maximum duration a route can be suppressed once it has been suppressed.

After enabling route dampening, the router never removes a route from the BGP table. A route that has been withdrawn by a BGP neighbor can still be seen in the BGP table and is marked with an "h" (history state).

A penalty is always applied to a path and not a prefix. If one of the paths is flapping, it does not mean that the destination is flapping.

 Copyright © 2003, Cisco Systems, Inc.

Practice

Q1) What are two things that happen to an EBGP route that has become unreachable when BGP route dampening is used? (Choose two.)

A) It is removed from the IP routing table.

B) It is removed from the BGP table.

C) It will remain in the IP routing table as long as its penalty remains greater than 50 percent of the reuse limit.

D) It is kept in the BGP table and marked as a history entry.

Q2) What type of routes can be dampened?

A) IGP

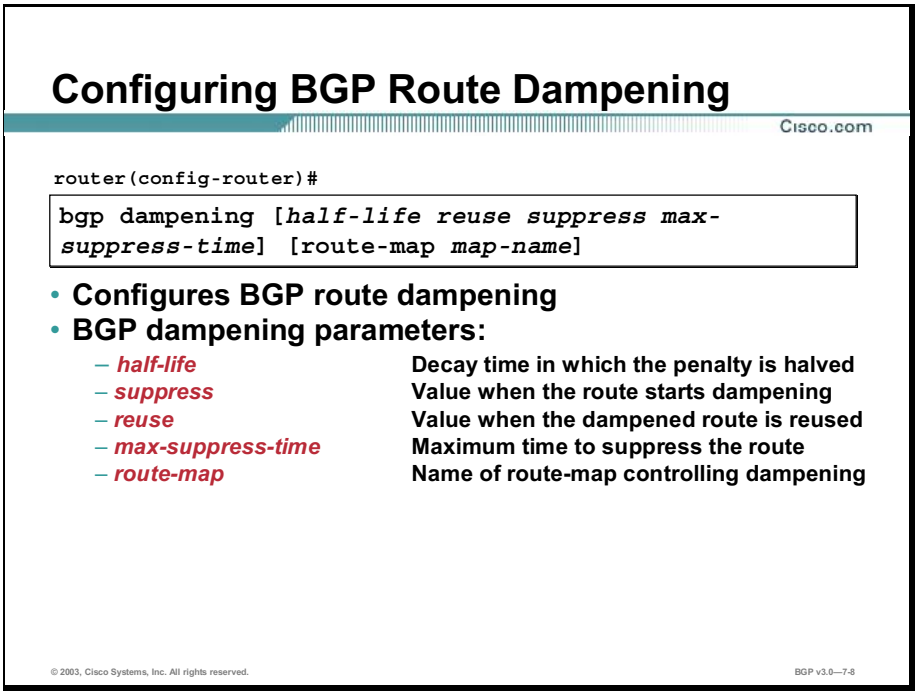
B) IBGP

C) EBGP

D) locally advertised routes

Configuring BGP Route Dampening

This topic identifies the Cisco IOS commands required to configure BGP route dampening.


Configuring BGP Route Dampening
Cisco.com
router(config-router)#
bgp dampening [half-life reuse suppress max-suppress-time] [route-map map-name]
• Configures BGP route dampening
• BGP dampening parameters:
– half-life
– suppress
– reuse
– max-suppress-time
– route-map
Decay time in which the penalty is halved
Value when the route starts dampening
Value when the dampened route is reused
Maximum time to suppress the route
Name of route-map controlling dampening
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-8

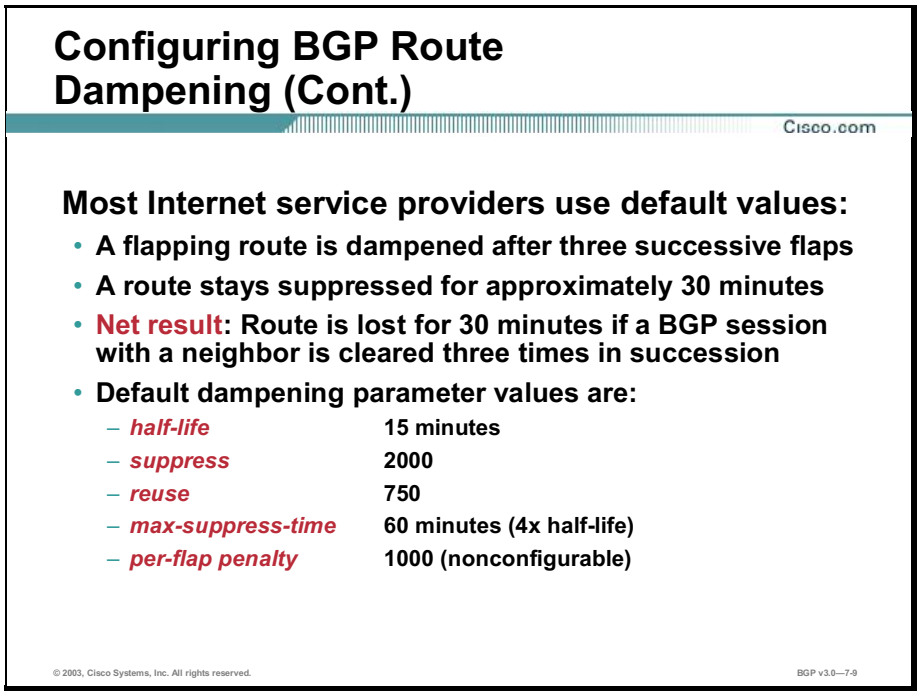
To enable route dampening, use the **bgp dampening** command. Optionally, you can change the default settings of the route dampening parameters.

Route flap dampening requires the following parameters:

- *half-life*: The half-life is the time needed for the penalty to halve (default is 15 minutes).
- *suppress*: When a route has more penalty points than the suppress limit, the route is suppressed (default is 2000).
- *reuse*: After the flapping has stopped and the penalty for a route has fallen below the reuse limit, the route is unsuppressed (default is 750).
- *max-suppress-time*: No route can be suppressed longer than the max-suppress-time minutes (default is 1 hour; maximum is 255 minutes).

You can specify the four route flap dampening parameters directly with the **bgp dampening** command. Alternatively, you can create a route-map that specifies different dampening parameters for different sets of routes and apply the route-map with the **bgp dampening route-map** command.

 Copyright © 2003, Cisco Systems, Inc.


Configuring BGP Route Dampening (Cont.)
Cisco.com
Most Internet service providers use default values:
• A flapping route is dampened after three successive flaps
• A route stays suppressed for approximately 30 minutes
• Net result: Route is lost for 30 minutes if a BGP session with a neighbor is cleared three times in succession
• Default dampening parameter values are:
– half-life 15 minutes
– suppress 2000
– reuse 750
– max-suppress-time 60 minutes (4x half-life)
– per-flap penalty 1000 (nonconfigurable)
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-9

This sample calculation shows how long a route that flaps three times is suppressed with the default values of the Cisco IOS route-dampening parameters. Each time a route flaps it accumulates another 1000 points. After the third flap, the route has almost 3000 points. Remember that the penalty is gradually decreased using a decaying algorithm, causing a reduction in the number of points the route accumulates. It takes 15 minutes for the penalty to drop below 1500 (provided there are no further flaps) and another 15 minutes to drop below the reuse limit (750).

Many service providers change the default parameters to allow a maximum suppress time of several hours.

Note	Using the **clear ip bgp *** command is regarded as a flap to neighboring autonomous systems. Using this command several times may cause neighboring autonomous systems to suppress prefixes for some time even if there is nothing wrong with the route.

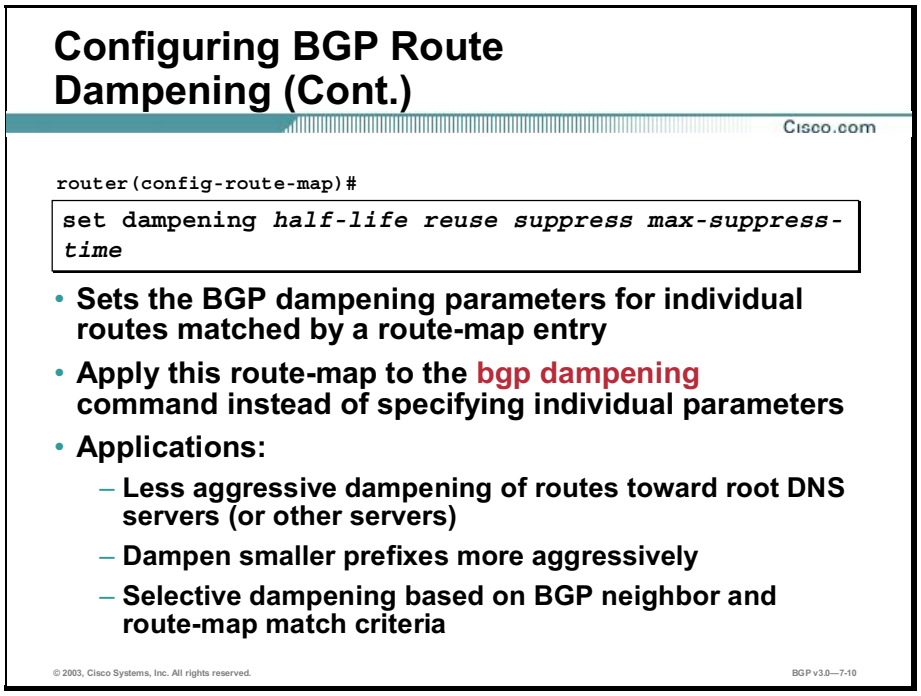
Note	Using the **clear ip bgp * [soft] [in|out]** command is not regarded as a flap and should be used instead of **clear ip bgp *** for clearing the neighbor relationships.

The following default dampening parameter values are used if they are not specified by configuration:

- *half-life* 15 minutes
- *suppress* 2000
- *reuse* 750
- *max-suppress-time* 60 minutes
- *per-flap penalty* 1000

You can change all default values if you specify them in the optional parameters of the **bgp dampening** command. The per-flap penalty is the only value that is not configurable.

 Copyright © 2003, Cisco Systems, Inc.


Configuring BGP Route Dampening (Cont.)
Cisco.com
router(config-route-map)#
set dampening half-life reuse suppress max-suppress-time
• Sets the BGP dampening parameters for individual routes matched by a route-map entry
• Apply this route-map to the bgp dampening command instead of specifying individual parameters
• Applications:
– Less aggressive dampening of routes toward root DNS servers (or other servers)
– Dampen smaller prefixes more aggressively
– Selective dampening based on BGP neighbor and route-map match criteria
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-10

Many service providers prefer to implement selective dampening. Larger prefixes are usually less likely to flap and should not be penalized as aggressively as smaller prefixes that populate the majority of the BGP table.

You can use a route-map in combination with a prefix-list to match on prefix length and to set different route-dampening parameters for larger prefixes than for smaller ones. A practical service provider policy is to use a route-map to exclude root Domain Name System (DNS) servers from dampening altogether.

You can then attach the route-map to the BGP route dampening process with the **bgp dampening route-map** command.

Practice

Q1) What are two methods of enabling route dampening on a Cisco router? (Choose two.)

A) globally, by enabling route dampening in global router configuration mode

B) globally, by enabling route dampening under the BGP routing process

C) on specific routes by enabling route dampening on a specific interface

D) by using a route-map in the BGP process to apply route dampening to specific routes

Q2) Given a reuse of 1000 and a suppress of 3000, how should you configure the *half-life* parameter if a route flapping four times is reusable after 20 minutes?

A) 5 min

B) 10 min

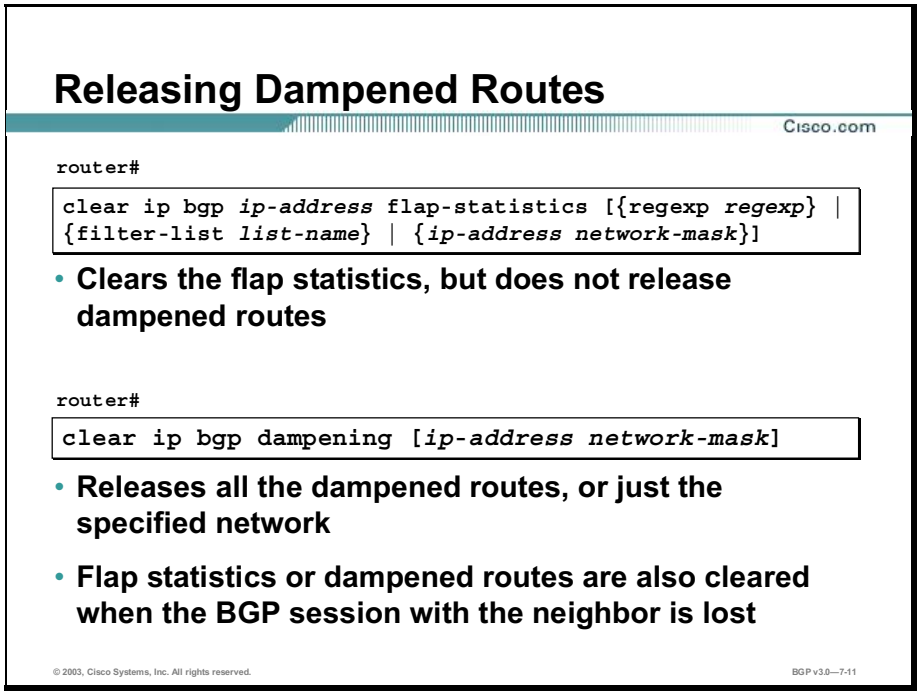
C) 15 min

D) 20 min

 Copyright © 2003, Cisco Systems, Inc.

Releasing Dampened Routes

This topic identifies the Cisco IOS commands used to release dampened routes.


Releasing Dampened Routes
Cisco.com
router#
clear ip bgp ip-address flap-statistics [{regexp regexp} | {filter-list list-name} | {ip-address network-mask}]
• Clears the flap statistics, but does not release dampened routes
router#
clear ip bgp dampening [ip-address network-mask]
• Releases all the dampened routes, or just the specified network
• Flap statistics or dampened routes are also cleared when the BGP session with the neighbor is lost
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-11

There are two timers calculated for every route when it flaps:

- **Time to forget:** The time that it takes before all flap history is deleted. Using the **clear ip bgp flap-statistics** command deletes all penalty information, but it does not release the dampened paths.

- **Time to reuse:** The time that it takes before a route can be considered for best-path processing. Using the **clear ip bgp dampening** command resets this timer for all networks or for specified networks so that they are no longer suppressed. The flap statistics, however, are still kept, and the next flap will cause the previously dampened paths to be suppressed again.

clear ip bgp flap-statistics

To clear BGP flap statistics, use the **clear ip bgp flap-statistics** privileged EXEC command.

clear ip bgp ip-*address* **flap-statistics** [{**regexp** *regexp*} | {**filter-list** *list-name*} | {*ip-address network-mask*}]

Syntax description

regexp *regexp*	(Optional) Clears flap statistics for all the paths that match the regular expression.

filter-list *list-name*	(Optional) Clears flap statistics for all the paths that pass the access-list.
ip-address	(Optional) Clears flap statistics for a single entry at this IP address. If this argument is placed before **flap-statistics**, the router clears flap statistics for all paths from the neighbor at this address.
network-mask	(Optional) Network mask applied to the *ip-address* argument.

clear ip bgp dampening

To clear BGP route dampening information and unsuppress the suppressed routes, use the **clear ip bgp dampening** privileged EXEC command.

clear ip bgp dampening [*ip-address network-mask*]

Syntax description

ip-address	(Optional) IP address of the network from which to clear dampening information
network-mask	(Optional) Network mask applied to the *ip-address* argument

Practice

Q1) A network has been suppressed due to route dampening. What are two things that you can do on the router to unsuppress the dampened route? (Choose two.)

A) Clear the flap statistics for the dampened prefix.

B) Issue a **clear ip bgp dampening** command on the router.

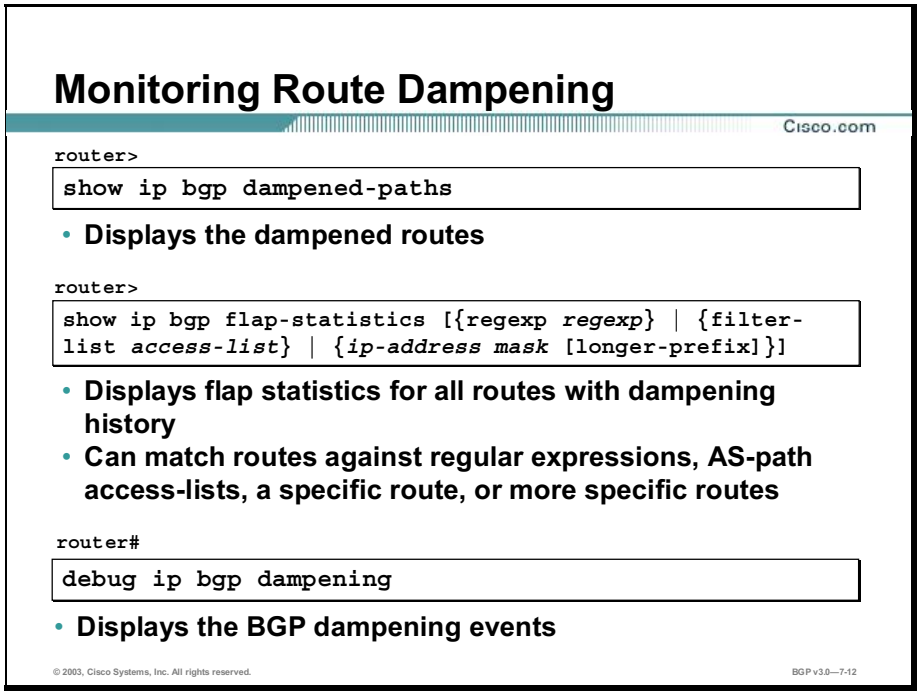
C) Wait for the route penalty to fall below the reuse limit.

D) Reset the BGP process on the neighbor advertising the route.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring Route Dampening

This topic identifies the Cisco IOS commands required to monitor BGP route dampening.


Monitoring Route Dampening
Cisco.com
router>
show ip bgp dampened-paths
• Displays the dampened routes
router>
show ip bgp flap-statistics [{regexp regexp} | {filter-list access-list} | {ip-address mask [longer-prefix]}]
• Displays flap statistics for all routes with dampening history
• Can match routes against regular expressions, AS-path access-lists, a specific route, or more specific routes
router#
debug ip bgp dampening
• Displays the BGP dampening events
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-12

The penalty placed on a network is decayed until the reuse limit is reached, upon which the route is once again advertised. Every time a route flap occurs, the penalty is recalculated. In the router, the penalty is encoded as the time it takes for the penalty to decay below the reuse limit. At half of the reuse limit, the dampening information for the route to the network is removed.

Use the **show ip bgp dampened-paths** command to list all routes that are currently suppressed due to dampening.

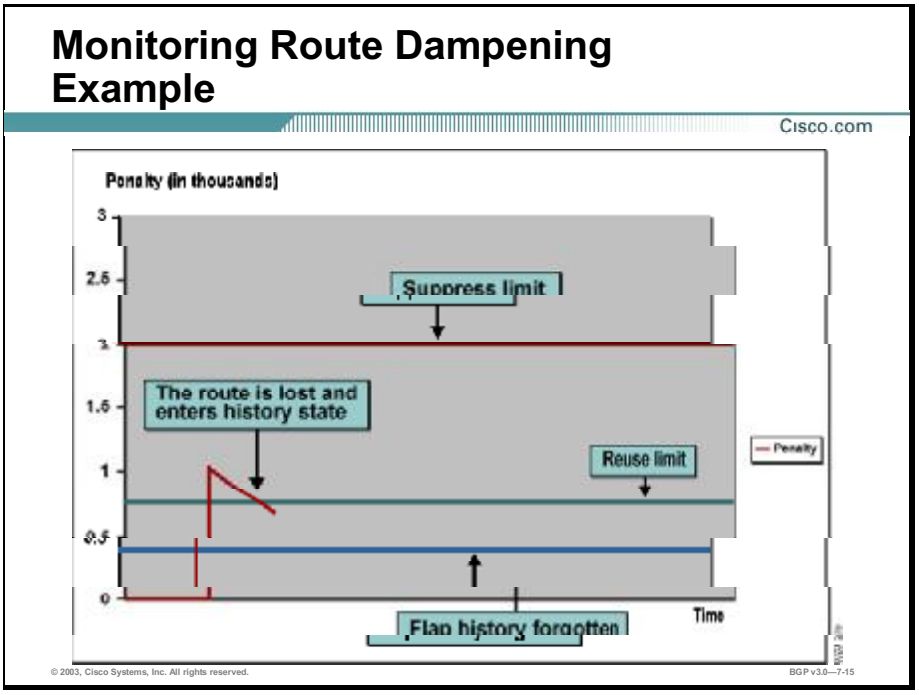
Use the **show ip bgp flap-statistics** command to list all routes that have a penalty still above the time-to-forget limit. You can also use this command in combination with regular expressions and filter lists.

The **show ip bgp flap-statistics prefix** command displays detailed dampening information about a specific network.

The **show ip bgp flap-statistics prefix mask longer-prefix** command displays dampening information about a specific network and its subnets.

Use the **debug ip bgp dampening** command to display BGP dampening events as they occur in real time.

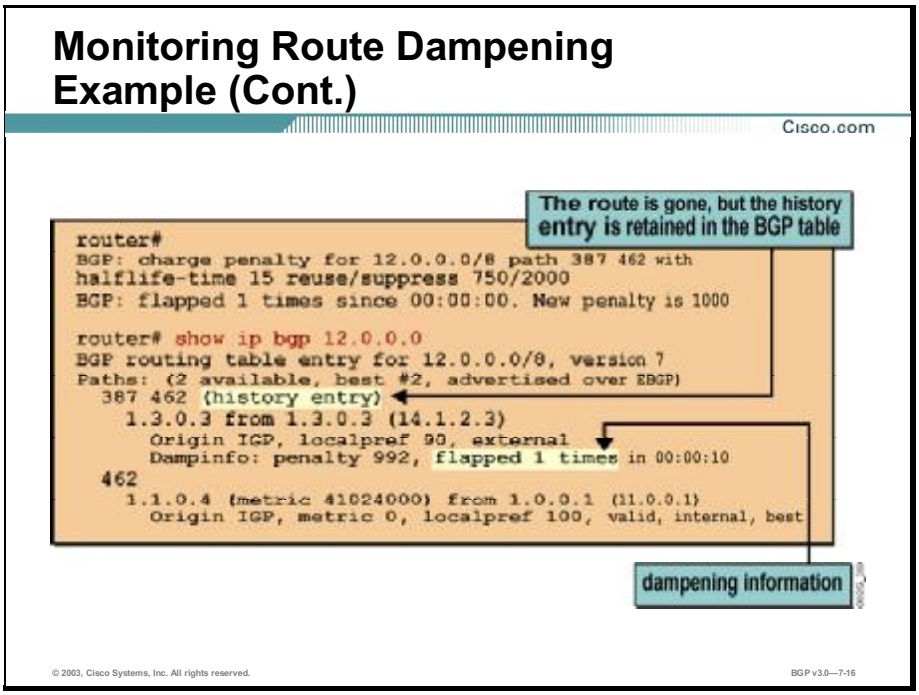
Example



The example here shows how, after the first flap (a route becomes unreachable), the router withdraws the route but keeps it in its own database to keep track of the penalty. The route enters the history state.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring Route Dampening Example (Cont.)

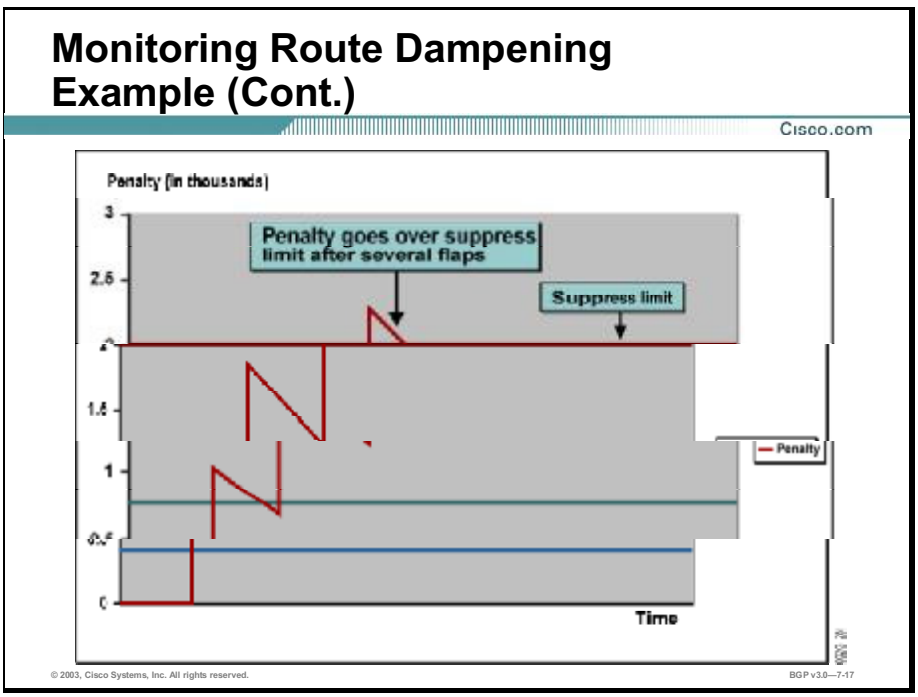

Cisco.com
The route is gone, but the history entry is retained in the BGP table
router#
BGP: charge penalty for 12.0.0.0/8 path 387 462 with halflife-time 15 reuse/suppress 750/2000
BGP: flapped 1 times since 00:00:00. New penalty is 1000
router# show ip bgp 12.0.0.0
BGP routing table entry for 12.0.0.0/8, version 7
Paths: (2 available, best #2, advertised over EBGP)
387 462 (history entry)
1.3.0.3 from 1.3.0.3 (14.1.2.3)
Origin IGP, localpref 90, external
Dampinfo: penalty 992, flapped 1 times in 00:00:10
462
1.1.0.4 (metric 41024000) from 1.0.0.1 (11.0.0.1)
Origin IGP, metric 0, localpref 100, valid, internal, best
dampening information
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-16

Using **show ip bgp** displays suppressed prefixes with state “h”.

Using **show ip bgp prefix** displays suppressed prefixes marked with “history entry”.

If a prefix is in the history state, it means that it is currently unreachable but the information is kept in the BGP table to keep track of the penalty.

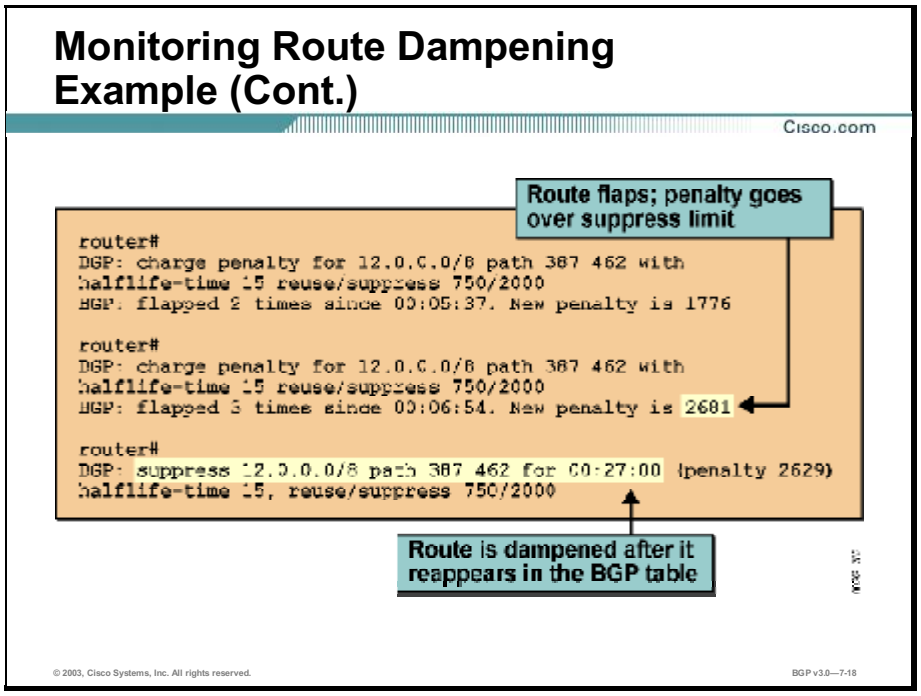
Monitoring Route Dampening Example (Cont.)



The example here shows how after the third flap, the penalty of the route exceeds the suppress limit, and the route could be suppressed. When the route exceeds the suppress limit, one of two things could happen:

- The router will put the route in the history state if it is currently unreachable.
- The router will suppress the route if it is currently reachable.

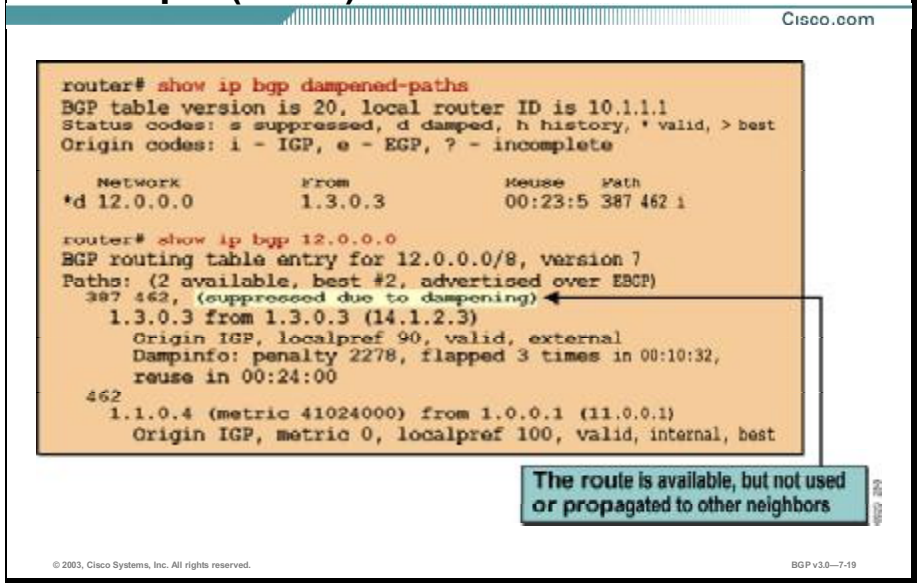
 Copyright © 2003, Cisco Systems, Inc.


Monitoring Route Dampening Example (Cont.)
Cisco.com
Route flaps; penalty goes over suppress limit
router#
DGP: charge penalty for 12.0.0.0/8 path 387 462 with
halflife-time 15 reuse/suppress 750/2000
BGP: flapped 2 times since 00:05:37. New penalty is 1776
router#
DGP: charge penalty for 12.0.0.0/8 path 387 462 with
halflife-time 15 reuse/suppress 750/2000
BGP: flapped 3 times since 00:06:54. New penalty is 2681
router#
DGP: suppress 12.0.0.0/8 path 387 462 for 00:27:00 (penalty 2629)
halflife-time 15, reuse/suppress 750/2000
Route is dampened after it reappears in the BGP table
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-18

The figure displays the debugging messages that accompany the three route flaps illustrated in the previous figure. After the 12.0.0.0/8 prefix flaps for the third time, the router assigns a new penalty of 2681 to the route. Because the new penalty exceeds the suppress limit of 2000, the 12.0.0.0/8 prefix is suppressed by the router for 27 minutes (if it remains stable).

Using **show ip bgp** displays all suppressed prefixes that have their state marked as dampened ("d").

Monitoring Route Dampening Example (Cont.)

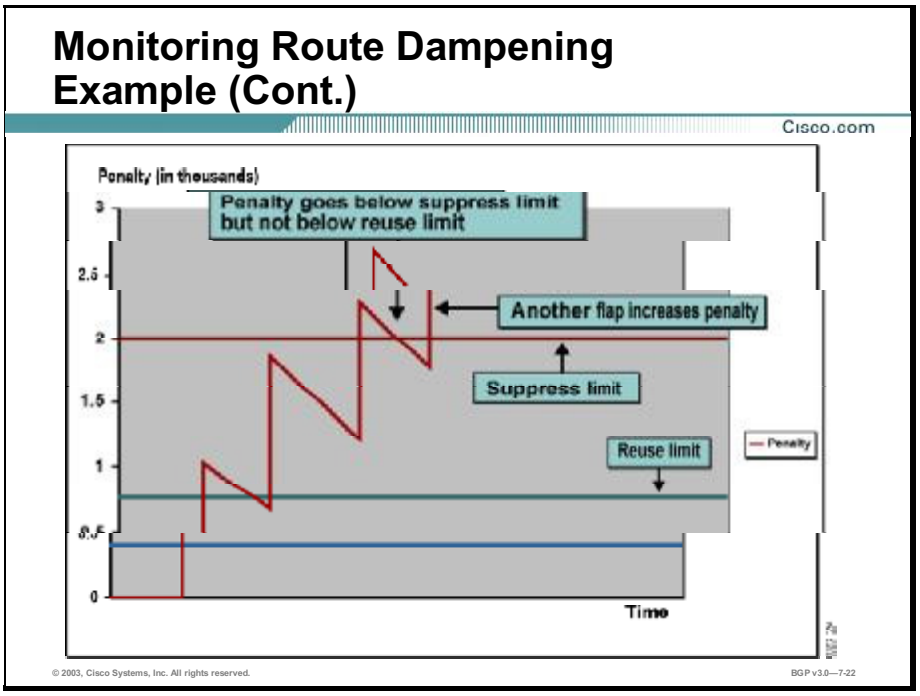

Cisco.com
router# show ip bgp dampened-paths
BGP table version is 20, local router ID is 10.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best
Origin codes: i - IGP, e - EGP, ? - incomplete
Network From Reuse Path
*d 12.0.0.0 1.3.0.3 00:23:5 387 462 i
router# show ip bgp 12.0.0.0
BGP routing table entry for 12.0.0.0/8, version 7
Paths: (2 available, best #2, advertised over EBGP)
387 462, (suppressed due to dampening)
1.3.0.3 from 1.3.0.3 (14.1.2.3)
Origin IGP, localpref 90, valid, external
Dampinfo: penalty 2278, flapped 3 times in 00:10:32,
reuse in 00:24:00
462
1.1.0.4 (metric 41024000) from 1.0.0.1 (11.0.0.1)
Origin IGP, metric 0, localpref 100, valid, internal, best
The route is available, but not used or propagated to other neighbors
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-19

The **show ip bgp dampened-paths** command is used to list all networks that are currently suppressed due to dampening.

The **show ip bgp prefix** command will display detailed information among other paths about suppressed prefixes. These prefixes are marked with the words “suppressed due to dampening.”

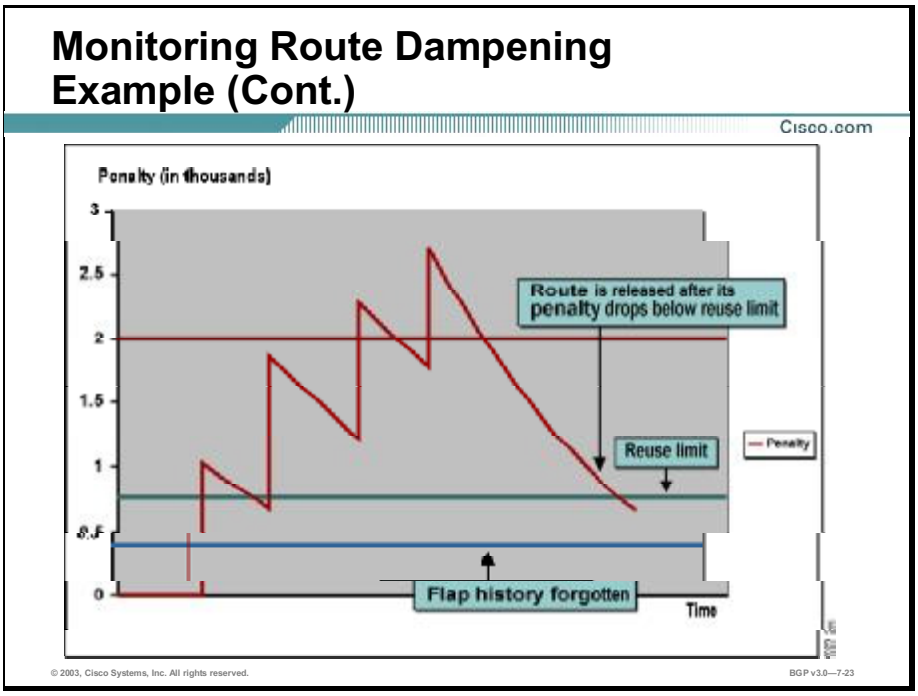
 Copyright © 2003, Cisco Systems, Inc.

Monitoring Route Dampening Example (Cont.)



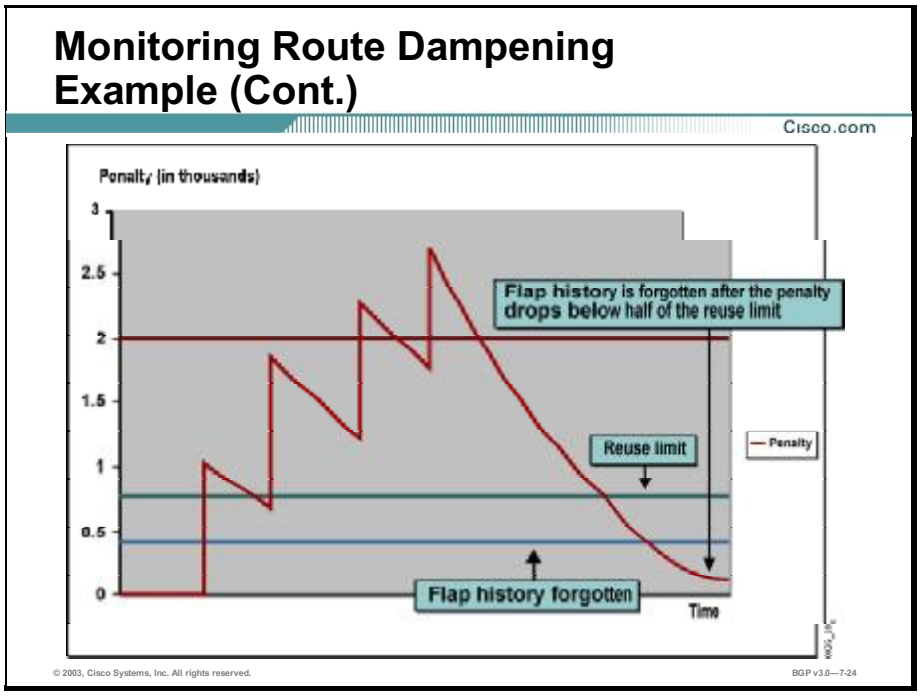
The penalty of the route is decreased following an exponential curve. After a while, the penalty drops below the suppress limit, but the route is not yet released—the route is released only after the penalty drops further below the reuse limit. In the example here, the route flaps again, further increasing the penalty.

Monitoring Route Dampening Example (Cont.)



In the example here, the route has stabilized, and the penalty gradually drops below the reuse limit. At that point in time, the BGP router releases the route, and it can now be selected as the best BGP path. If the released route is selected by the router as the best BGP path, it is also propagated to BGP neighbors and used for data forwarding.

 Copyright © 2003, Cisco Systems, Inc.



When the penalty associated with a route drops below half of the reuse limit, the penalty and the flap history associated with the route are cleared by the router.

Practice

Q1) What two things could happen to a BGP route that is penalized above the reuse limit, but its assigned penalty is under the suppress limit? (Choose two.)

A) The route is suppressed from BGP updates if it is reachable.

B) The route is marked as a history entry in the BGP table.

C) The route is withdrawn from the IP routing table.

D) The route will continue to be advertised.

Q2) When you are using default route-dampening parameter values, what happens to a suppressed route after its penalty decays for a time period equal to the half-life?

A) The route is released and available as a BGP candidate for best-path selection.

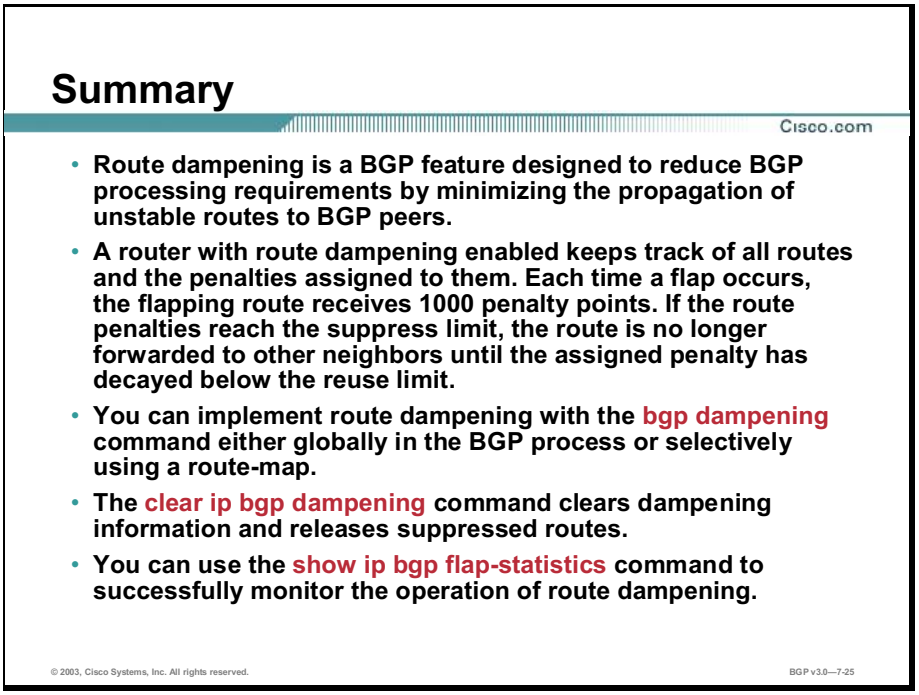
B) The route flap statistics are cleared, but the route will remain suppressed.

C) The route will remain suppressed for at least one more half-life.

D) The route will be released provided that no additional flaps have occurred.

Summary

This topic summarizes the key points discussed in this lesson.


Summary
Cisco.com
• Route dampening is a BGP feature designed to reduce BGP processing requirements by minimizing the propagation of unstable routes to BGP peers.
• A router with route dampening enabled keeps track of all routes and the penalties assigned to them. Each time a flap occurs, the flapping route receives 1000 penalty points. If the route penalties reach the suppress limit, the route is no longer forwarded to other neighbors until the assigned penalty has decayed below the reuse limit.
• You can implement route dampening with the bgp dampening command either globally in the BGP process or selectively using a route-map.
• The clear ip bgp dampening command clears dampening information and releases suppressed routes.
• You can use the show ip bgp flap-statistics command to successfully monitor the operation of route dampening.
© 2003, Cisco Systems, Inc. All rights reserved.
BGP v3.0—7-25

References

For additional information, refer to these resources:

- For more information on BGP route dampening and its function, refer to "BGP Case Studies Section 4" at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#flapdampen

- For more information on configuring BGP route dampening and its function, refer to "Configuring BGP" at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121cgcr/ip_c/ipcprt2/1cdbgp.htm - xtocid61

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise: BGP Route Dampening

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will use route dampening to minimize the impact of unstable routes.

After completing this exercise, you will be able to:

- Configure BGP route dampening
- Release dampened routes
- Monitor BGP route dampening

Copyright © 2003, Cisco Systems, Inc.

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
bgp dampening route-map *route-map*	Specify BGP dampening parameters using a route-map.
match community *c-acl*	Use this command in a route-map to match BGP communities by using a community-list.
set dampening *half-life reuse suppress max-suppress-time*	To set the BGP route-dampening factors
ip community-list *num* {**permit**|**deny**} *community*	Create a community access-list.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp flap-statistics	View prefixes that have previously flapped.
show ip bgp dampened-paths	View prefixes that are currently suppressed.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.
show ip bgp community [*community* [*community* …]] [**exact-match**]	Use this command to view BGP routes that have at least one BGP community attribute or those specified in the command.
show ip bgp community-list *c-list* [**exact-match**]	Use this command to view BGP routes that are permitted by the specified community-list.
debug ip bgp dampening	Use this command to see major events related to route dampening.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- Service providers usually use BGP route dampening to minimize the impact that route flaps in the Internet have on the stability of their networks. BGP route dampening also reduces the CPU load on the routers running BGP. In this exercise, you will minimize the impact of flapping routes on your AS and upstream autonomous systems by implementing graded BGP route dampening.

- Configure graded route dampening for routes received from the “Client” router according to the following design:

  - If a prefix contains a community attribute of x:300, then the prefix should not be dampened.

  - Prefixes longer than /23 should have a half-life of 20 minutes.

  - Prefixes between /9 and /22 should have a half-life of 15 minutes.

 Copyright © 2003, Cisco Systems, Inc.

- Prefixes between /0 and /8 should have a half-life of 10 minutes.
- Maximum suppression time should be changed to four times the half-life time for all prefixes.

- Figure 1 shows the connectivity that is established between your AS and the “Client” router.

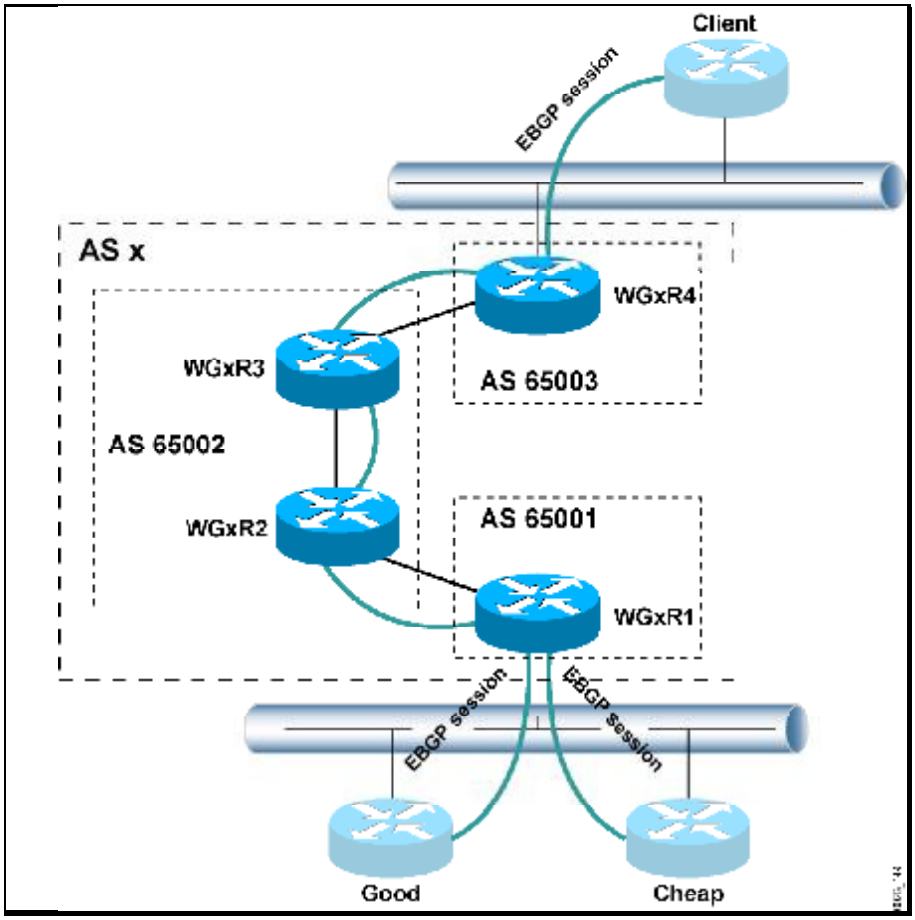


Figure 1: Route-dampening exercise topology

Exercise Procedure

Complete these steps:

Creating filters for the dampening route-map

Step 1 Create a community-list that matches prefixes carrying BGP community x:300 (where x is your real AS number, not your member-AS number).

Step 2 Create a prefix-list that matches all prefixes longer than /23.

Step 3 Create a prefix-list that matches all prefixes longer than /8.

Creating a route-map for route dampening

Step 4 Create a new route-map. The first statement of the new route-map should match prefixes carrying BGP community x:300. No route-dampening parameters need to be set.

Step 5 The second statement should match prefixes longer than /23. Set the half-life to 20 minutes and the max-suppress-time to 80 minutes, and use default values for other parameters (reuse 750, suppress 2000).

Step 6 The third statement should match prefixes longer than /8. Set the half-life to 15 minutes and the max-suppress-time to 60 minutes, and use default values for other parameters (reuse 750, suppress 2000).

Step 7 The last statement should match all other prefixes (**no match** command). Set the half-life to 10 minutes and the max-suppress-time to 40 minutes, and use default values for other parameters (reuse 750, suppress 2000).

Step 8 Enter BGP configuration mode and start BGP route dampening based on the created route-map with the **bgp dampening route-map** *route-map* command.

Exercise Verification

You have completed this exercise when you attain these results:

- Shutdown interface Loopback55 on router "Client" several times to create flaps. Use **show ip bgp flap-statistics** and **show ip bgp dampened-paths** to verify the operation of route dampening on your WGxR4 router.

```
WGxR4#sh ip bgp damp
BGP table version is 89, local router ID is 197.1.7.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          From             Reuse    Path
*d 99.0.0.0         192.168.21.99    00:25:20 99 i
*d 197.99.128.0/20  192.168.21.99    00:40:30 99 i


WGxR4#sh ip bgp flap
BGP table version is 89, local router ID is 197.1.7.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
```

 Copyright © 2003, Cisco Systems, Inc.

```
   Network          From            Flaps Duration Reuse    Path
*d 99.0.0.0         192.168.21.99   6     00:09:25 00:25:20 99
*d 197.99.128.0/20  192.168.21.99   6     00:09:25 00:40:20 99


WGxR4#sh ip bgp 99.0.0.0
BGP routing table entry for 99.0.0.0/8, version 80
Paths: (1 available, no best path)
  Not advertised to any peer
  99, (suppressed due to dampening)
    192.168.21.99 from 192.168.21.99 (197.99.111.1)
      Origin IGP, metric 0, localpref 100, valid, external
      Dampinfo: penalty 958, flapped 6 times in 00:31:20, reuse in
00:03:30


WGxR4#sh ip bgp
BGP table version is 89, local router ID is 197.1.7.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric LocPrf Weight Path
*> 10.0.0.0         192.168.20.233           0    100      0 (65002
65001) 20 i
*d 99.0.0.0         192.168.21.99            0             0 99 i
*> 128.20.0.0       192.168.20.20            0    100      0 (65002
65001) 20 i
*> 128.20.12.0/24   192.168.20.20            0    100      0 (65002
65001) 20 i
*> 128.22.0.0       192.168.20.20            0    100      0 (65002
65001) 20 i
*> 128.22.12.0/24   192.168.20.22            0    100      0 (65002
65001) 22 i
*> 128.26.0.0       192.168.20.22            0    100      0 (65002
65001) 22 26 i
*> 128.37.0.0       192.168.20.20            0    100      0 (65002
65001) 20 42 37 i
*> 128.42.0.0       192.168.20.20            0    100      0 (65002
65001) 20 42 i
*> 128.51.0.0       192.168.20.22            0    100      0 (65002
65001) 22 26 51 i
*> 128.213.0.0      192.168.20.20            0    100      0 (65002
65001) 20 213 i
*> 128.214.0.0      192.168.20.22            0    100      0 (65002
65001) 22 214 i
*> 192.20.11.0      192.168.20.20            0    100      0 (65002
65001) 20 i
*> 192.20.12.0/30   192.168.20.20            0    100      0 (65002
65001) 20 i
*> 192.22.11.0      192.168.20.20            0    100      0 (65002
65001) 20 i
*> 192.22.12.0/30   192.168.20.22            0    100      0 (65002
65001) 22 i
```

```
*> 192.26.11.0      192.168.20.22            0    100      0 (65002
65001) 22 26 i
*> 192.37.11.0      192.168.20.20            0    100      0 (65002
65001) 20 42 37 i
*> 192.42.11.0      192.168.20.20            0    100      0 (65002
65001) 20 42 i
*> 192.51.11.0      192.168.20.22            0    100      0 (65002
65001) 22 26 51 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
*> 192.213.11.0     192.168.20.20            0    100      0 (65002
65001) 20 213 i
*> 192.214.11.0     192.168.20.22            0    100      0 (65002
65001) 22 214 i
*> 197.1.1.0        197.1.1.1                0    100      0 (65002
65001) i
*> 197.1.2.0        197.1.2.1                0    100      0 (65002)
i
*> 197.1.3.0        197.1.2.1                0    100      0 (65002)
i
*> 197.1.4.0        197.1.4.1                0    100      0 (65002)
i
*> 197.1.5.0        197.1.4.1                0    100      0 (65002)
i
*> 197.1.6.0        0.0.0.0                  0         32768 i
*> 197.1.7.0        0.0.0.0                  0         32768 i
*> 197.1.8.0        197.1.1.1                0    100      0 (65002
65001) i
*> 197.99.1.0       192.168.21.99            0             0 99 i
*> 197.99.11.0      192.168.21.99            0             0 99 i
*> 197.99.12.0      192.168.21.99            0             0 99 i
*> 197.99.13.0      192.168.21.99            0             0 99 i
*> 197.99.20.0      192.168.21.99            0             0 99 i
*> 197.99.22.0      192.168.21.99            0             0 99 i
*> 197.99.80.0      192.168.21.99            0             0 99 i
*> 197.99.111.1/32  192.168.21.99            0             0 99 i
*> 197.99.120.0     192.168.21.99            0             0 99 i
*d 197.99.128.0/20  192.168.21.99            0             0 99 i
*> 200.20.0.0/16    192.168.20.20            0    100      0 (65002
65001) 20 i
*> 200.22.0.0/16    192.168.20.22            0    100      0 (65002
65001) 22 i
```

Answer these questions:

Q1) What is the purpose of route dampening?

Q2) Which routes are affected by route dampening?

 Copyright © 2003, Cisco Systems, Inc.

A

Answer Key

This document presents the correct answers and solutions for the course practices, assessments, and lab exercises.

Outline

This resource includes these topics:

- Practice Items
- Lesson Assessments
- Laboratory Exercises

 Copyright © 2003, Cisco Systems, Inc.

Practice Items

The practice items and solutions are listed here. The shaded answer options are the correct answers. This is a presentation of all course practice items.

Module 1: BGP Overview

Introduction to BGP

Practice: Interdomain Routing

Q1) What is an AS?

A) a set of Internet routers maintained by the IANA

B) a set of routing protocols designed to run inside a corporation or customer network

C) a set of protocols designed to implement routing policies

D) a collection of networks under a single technical administration

Q2) What are the three design goals of an interdomain routing protocol? (Choose three.)

A) secure routing information exchange

B) support for routing policies

C) ability to provide WAN connectivity from remote customer sites

D) ability to scale to an extremely large number of routes

Practice: BGP Characteristics

Q1) Which transport mechanism is used to exchange BGP routing updates?

A) UDP

B) TCP

C) SDP

D) LDAP

Q2) To achieve reliability and scalability, what BGP design tradeoff was made?

A) Route authentication is not supported.

B) Convergence is slower due to batched updates and long keepalive timers.

C) BGP supports only specific IGPs.

D) BGP has to perform periodic updates to synchronize route tables.

Q3) What are BGP keepalive messages used for?

A) to form the initial adjacency between BGP peers

B) to calculate the best path by determining the round-trip delay

C) to detect a neighbor when the BGP session is in the idle state

D) to acknowledge the receipt of BGP route updates from the BGP neighbor

Practice: Single-Homed Customers

Q1) What are the two reasons to use BGP between a single-homed customer and a service provider? (Choose two.)

A) if multiple connections to the same provider will be used

B) if the customer is using RFC 1918 IP addressing

C) as a preparation step for a future upgrade to multiple providers

D) if the customer has a large WAN with a single Internet connection

 Copyright © 2003, Cisco Systems, Inc.

Practice: Multihomed Customers

Q1) Why should the multihomed customer avoid forwarding any routing information received from one ISP to the other?

A) to guarantee that routing loops will not occur

B) to avoid the possibility that the customer could become a transit backbone between the two providers

C) to prevent route leaking of the customer IGP routes into the Internet

D) to prevent denial-of-service attacks from the Internet

Practice: Transit Autonomous Systems

Q1) What is the main function of a transit AS in BGP?

A) to facilitate customer connectivity to the Internet

B) to forward routing information learned from one AS into another AS

C) to direct routing policy in neighboring autonomous systems

D) to allow non-BGP configured customer routers to connect to the Internet

Practice: BGP Limitations

Q1) What are BGP forwarding decisions based upon?

A) source IP address

B) destination IP address

C) next-hop attribute

D) remote AS policies

Copyright © 2003, Cisco Systems, Inc.

BGP Session Establishment

Practice: BGP Neighbor Discovery

Q1) How are BGP neighbors discovered?

A) Through the use of broadcast packets.

B) Through the use of IP multicast packets.

C) Through the use of a hello protocol.

D) Neighbors are not discovered but instead must be manually configured.

Q2) What protocol and port are used by BGP for session establishment? (Choose two.)

A) TCP

B) UCP

C) 53

D) 161

E) 179

Q3) What is the state of a BGP router prior to forming a BGP neighbor session?

A) Down

B) Idle

C) Not Connected

D) Unestablished

Practice: Establishing a BGP Session

Q1) Which three steps must be taken before a BGP session is established? (Choose three.)

A) Neighbors must be manually configured as peers.

B) A TCP session between the two routers must be established.

C) The network topology database must be exchanged.

D) BGP Open messages must be exchanged and verified.

 Copyright © 2003, Cisco Systems, Inc.

Q2) What three parameters are carried in a BGP Open message? (Choose three.)

A) BGP version number

B) local AS number

C) BGP hold timer

D) current neighbor state

E) network link-state database

Practice: BGP Keepalives

Q1) What is the purpose of the keepalive in BGP?

A) to provide the clock used for packet transmission

B) to synchronize the TCP neighbor sessions

C) to provide reliable transmission when you are using BGP between neighbors

D) to verify the presence of a configured BGP neighbor

Practice: MD5 Authentication

Q1) What are the benefits of using MD5?

A) It encrypts routing updates between peers.

B) It provides for payload compression.

C) It ensures a reliable connection with retransmission of lost packets.

D) It provides peer authentication to prevent DoS attacks.

BGP Route Processing

Practice: Receiving Routing Updates

Q1) What three BGP attributes will always be present in received BGP routing updates? (Choose three.)

A) next-hop

B) local preference

C) origin

D) AS-path

Practice: Building the BGP Table

Q1) Which incoming BGP updates are stored in the BGP table?

A) only the routes selected as the best

B) only the routes that are learned from external BGP peers

C) all routes received from BGP neighbors

D) all routes received from Interior Gateway Protocols

Q2) What does the router use in the output of the **show ip bgp** command to indicate the best route?

A) i

B) *

C) >

D) B

 Copyright © 2003, Cisco Systems, Inc.

Practice: BGP Route Selection Criteria

Q1) With no BGP attributes modified, which criterion will most likely determine route selection?

A) weight

B) local preference

C) MED

D) AS-path length

Practice: BGP Route Propagation

Q1) Which BGP routes are announced to other BGP neighbors?

A) All learned routes are sent to all BGP neighbors.

B) Only the routes learned from IGPs are sent to other neighbors.

C) Only the routes selected as the best are announced to other neighbors.

D) All routes learned from EBGP neighbors are sent to all other BGP peers.

Practice: Building the IP Routing Table

Q1) What routing characteristic does a router use to select the best route when multiple sources exist for the same route?

A) cost

B) AD

C) weight

D) metric

Copyright © 2003, Cisco Systems, Inc.

Practice: Advertising Local Networks

Q1) What are two methods for advertising local routes in BGP? (Choose two.)

A) announcing networks using the **network** command

B) sending and receiving routing updates using the **neighbor** command

C) redistribution of routes learned by another routing protocol

D) automatic redistribution of local routes to BGP peers

Practice: Automatic Summarization

Q1) Which combination of commands would result in inserting only more specific routes into the BGP table?

A) classful **network** command with automatic summarization enabled

B) redistribution into BGP with automatic summarization enabled

C) classful **network** command with automatic summarization disabled

D) redistribution into BGP with automatic summarization disabled

 Copyright © 2003, Cisco Systems, Inc.

Basic BGP Configuration

Practice: BGP Routing Process

Q1) How many BGP processes can be active in a router?

A) 1

B) 16

C) 256

D) depends on the configured memory in the router

Practice: Configuring External Neighbors

Q1) Which two parameters must you configure for a BGP neighbor? (Choose two.)

A) IP address

B) description

C) AS number

D) password

Practice: Configuring BGP Timers

Q1) What are the default values assigned to the keepalive and holdtime timers in BGP?

A) keepalive: 30 sec, holdtime: 90 sec

B) keepalive: 30 sec, holdtime: 120 sec

C) keepalive: 60 sec, holdtime: 180 sec

D) keepalive: 120 sec, holdtime: 270 sec

Copyright © 2003, Cisco Systems, Inc.

Practice: Configuring MD5 Authentication

Q1) In a BGP configuration using MD5 authentication, how must you configure passwords?

A) All routers in the AS must have the same password.

B) All neighbors on a single router must share the same password.

C) Neighbor routers must share the same password.

D) The hashing algorithm requires unique MD5 passwords.

Practice: Announcing Networks in BGP

Q1) How is the **network** command in BGP different from other routing protocols?

A) The BGP **network** command signifies what router interfaces should run BGP.

B) The **network** command in BGP contains a mask.

C) The **network** command indicates what routes should be injected into the BGP table.

D) The **network** command is the same for all IP routing protocols.

Practice: Redistributing Routes into BGP

Q1) What is a potential problem caused by redistributing all routes from the IGP to BGP?

A) Private addresses (RFC 1918) or other networks that the local AS does not wish to advertise will be propagated into BGP.

B) Summary prefixes will be announced, creating routing “black holes.”

C) The BGP neighbor will be unable to apply incoming route filters.

D) The BGP routing protocol process will become unstable due to the faster convergence requirements of the IGP routes.

 Copyright © 2003, Cisco Systems, Inc.

Q2) What is a key advantage of using route-maps when redistributing IGP routes into BGP?

A) The weight attribute can be set.

B) Redistribution of routes can be matched to an access-list.

C) The direction of the redistribution can be controlled.

D) The origin attribute can be set to IGP.

Practice: Configuring Classless BGP

Q1) In BGP prefix notation, what does the number following the slash (/) indicate?

A) the number of bits in the subnet mask being set to 1

B) the number of subnet bits following the natural mask

C) the number of bits in the host portion of an IP address

D) the mask to use in the octet of subnetting

Q2) If you use the **network mask** option with the BGP **network** command, what must be present for the aggregate to appear in the BGP routing table?

A) a loopback interface on the router in the same network as the summary

B) an exact match of the network number and mask in the IP routing table

C) specification of the **summary-only** option

D) a default route to null 0 on the router

Practice: Aggregating BGP Networks

Q1) What is the purpose of the **summary-only** option when you are configuring BGP route aggregation?

A) to explicitly include summary routes in route aggregates

B) to summarize BGP routes to their classful boundaries

C) to suppress the advertising of routes more specific than the aggregate address

D) to prevent summary routes from being advertised within an AS

Copyright © 2003, Cisco Systems, Inc.

Q2) What are two potential drawbacks to using route summarization? (Choose two.)

A) causes loss of granular routing information

B) reduces the size of the routing table

C) prohibits redistribution of summarized routes

D) can cause suboptimal routing to occur

Practice: Multihomed Customer Problem

Q1) When a customer is connected to multiple providers, how can aggregation cause suboptimal path selection?

A) Aggregation causes all traffic to flow in the direction of the aggregated route.

B) Aggregation causes all attributes of the route to be lost, preventing neighboring BGP routers from completing path selection based on BGP attributes.

C) If both providers are not aggregating routes, the longest matching prefix rule will cause traffic to flow toward the nonaggregated route.

D) BGP will attempt to load-balance between multiple paths to the destination, forcing each session over a different network path.

 Copyright © 2003, Cisco Systems, Inc.

Monitoring and Troubleshooting BGP

Practice: Monitoring Overall BGP Routing

Q1) What three key pieces of information are contained in the output of the **show ip bgp summary** command? (Choose three.)

A) BGP memory use

B) BGP neighbors

C) BGP route table

D) BGP neighbor connection state

Practice: Monitoring BGP Neighbors

Q1) Which command do you use to display detailed BGP neighbor information?

A) **show ip bgp summary**

B) **show ip bgp**

C) **show ip bgp neighbors** *address*

D) **show ip bgp detail**

Practice: Monitoring the BGP Table

Q1) How does the output from the **show ip bgp** command indicate which route to a specific destination is selected as the best?

A) The route selected as the best is enclosed in parentheses: "( )".

B) The route selected as the best is marked with the character ">".

C) The route selected as best is marked with an asterisk: "*".

D) Only the best route is inserted into the BGP table.

Practice: Debugging BGP

Q1) What debug command should you enable to troubleshoot BGP session startup issues where the TCP connection never succeeds?

A) **ip bgp updates**

B) **ip packets**

C) **ip bgp keepalives**

D) **ip tcp transactions**

Q2) What two steps can you take to limit the amount of debug output when debugging BGP updates? (Choose two.)

A) specify the incoming interface

B) limit the debug to a specific BGP neighbor

C) limit the debug to specific networks using ACLs

D) increase the keepalive timer to reduce BGP update traffic

Practice: BGP Session Startup Problems

Q1) What are three common BGP session startup issues? (Choose three.)

A) The BGP neighbor never becomes active.

B) The BGP neighbor state oscillates between idle and active.

C) The BGP session is established, but the neighbor state is idle.

D) The BGP neighbor is active, but the BGP session is not established.

Practice: BGP Neighbor Not Reachable

Q1) What is the most common reason for a BGP session not leaving the Idle state?

A) The TCP port for the connection is not configured.

B) The external neighbor is not directly connected.

C) The TCP SYN packet is answered with an RST packet.

D) The neighbors have been configured with the same AS number.

 Copyright © 2003, Cisco Systems, Inc.

Q2) What can you use to establish a BGP session to external neighbors that are not directly connected?

A) IBGP

B) an IGP

C) static routes

D) **ebgp-multihop**

Practice: BGP Neighbor Not Configured

Q1) What will result from attempting to open a BGP connection with a neighbor that has not been properly configured for BGP?

A) The BGP session will remain in the Idle state.

B) The neighbor session will be established, and the session startup parameters will be negotiated over the TCP session.

C) The BGP session is immediately terminated with a TCP RST packet.

D) The BGP session will become “stuck in Active state.”

Practice: BGP AS Number Mismatch

Q1) What symptom usually indicates an AS number mismatch?

A) The BGP session is established and immediately closed.

B) A TCP RST packet is sent.

C) The TCP session goes into synsent state.

D) The BGP session remains in Idle state.

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

Module 2: Route Selection Using Policy Controls

Multihomed BGP Networks

Practice: Business Requirements of Multihomed BGP Networks

Q1) What are two reasons why a customer would want to connect to two ISPs? (Choose two.)

A) to expand capacity for Internet traffic

B) to better protect confidential information as it travels through the Internet

C) to provide redundancy to mission-critical services offered over the Internet

D) to efficiently route Internet traffic to two different divisions within the company

Practice: Technical Requirements of Multihomed BGP Networks

Q1) What are the two technical requirements for multihomed customers? (Choose two.)

A) The ISPs must assign a range of IP network numbers to the customer.

B) The customer network must exchange BGP information with each ISP network.

C) In most cases, the customer must have its own public AS number.

D) The customer network must not use AS-path filters.

Q2) Which routing protocol needs to be deployed between multihomed customers and the ISP?

A) any interior gateway protocol

B) static routes

C) default routes

D) BGP

Copyright © 2003, Cisco Systems, Inc.

Practice: BGP Route Selection Without Policies

Q1) In the absence of applied policy, what two factors will most likely influence BGP route selection? (Choose two.)

A) weight

B) AS path

C) origin

D) router-ID

Practice: Multihomed Customer Routing Policies

Q1) List three potential customer routing policies. (Choose three.)

A) One service provider is designated as primary, and the other is a backup.

B) Traffic is load-balanced across both ISP networks.

C) Traffic toward a specific destination goes through only one of the ISPs.

D) Traffic to direct customers of the ISPs goes direct; all other traffic goes through the primary ISP.

Practice: Influencing BGP Route Selection

Q1) Why do you need to influence BGP route selection rules?

A) because the default BGP route selection does not always result in optimum routing

B) because BGP route selection favors the path with the highest bandwidth, which is not always the best path

C) because the ISP will select the route if you do not influence BGP route selection

D) because BGP route selection often favors less stable routes

 Copyright © 2003, Cisco Systems, Inc.

Practice: BGP Filters

Q1) What two potential multihomed network issues can be prevented with IP prefix filters? (Choose two.)

A) the propagation of private AS numbers

B) the propagation of private addresses used in the network

C) the propagation of unreachable next-hop addresses

D) the propagation of more specific prefixes from an address range

AS-Path Filters

Practice: AS-Path Filtering Scenarios

Q1) Which three of the following situations are appropriate applications of AS-path filters? (Choose three.)

A) to ensure only locally originated routes are announced

B) to limit routes advertised from IBGP neighbors

C) to select a subset of all routes based on their originating AS

D) to limit neighbor route updates to specific AS-originated routes

Practice: AS-Path Regular Expressions

Q1) Which AS path is not matched by the regular expression “27”?

A) 100 27

B) 27 64 128

C) 10 12 182 77

D) 10 12 182 77 27 71

Practice: String Matching

Q1) Which AS path is matched by the regular expression “72$”?

A) 213 72 218 31 727

B) 27 317 271 50 72

C) 315 27 723 19 91

D) 72 591 368 20 87

 Copyright © 2003, Cisco Systems, Inc.

Q2) Which AS path is matched by the regular expression "$27_"?

A) 213 72 218 31 727

B) 27 317 271 50 72

C) 315 27 723 19 91

D) 72 591 368 20 27

Q3) What is the difference between the regular expressions "_100_" and "_100$"?

A) The first expression refers to routes that have the substring "100" in their AS paths; the second expression refers only to routes that are directly connected to AS 100.

B) The first expression refers to routes that have the substring "100" in their AS paths; the second expression refers only to routes that originated in AS 100.

C) The first expression refers to routes that go through AS 100; the second expression refers to routes that originated in AS 100.

D) The first expression refers to routes that are directly connected to AS 100; the second expression refers to routes that originated in AS 100.

Q4) What does the regular expression "^100$" match?

A) routes that originated in AS 100

B) routes that go through AS 100

C) routes that contain "100" in the AS path

D) routes that originate in the neighboring AS 100

Q5) How do you match AS paths that contain exactly two single-digit AS-numbers?

A) Use the expression "**".

B) Use the expression "..".

C) Use the expression "[0-9]_[0-9]".

D) Use the expression "^[0-9]_[0-9]$".

Practice: Applying AS-Path Filters

Q1) What happens to outgoing routes that are not selected by matching a configured AS-path filter?

A) Nonselected routes are removed from the IP routing table.

B) Nonselected routes are removed from the BGP table.

C) Nonselected routes are sent to the neighbor with a poisoned metric.

D) Nonselected routes are used by the local router only.

Practice: Configuring BGP AS-Path Filters

Q1) What three steps are required to apply a new inbound routing policy to a neighbor? (Choose three.)

A) Define an AS-path access-list.

B) Attach the AS-path filter to inbound or outbound updates for a specific BGP neighbor.

C) Send incoming and/or outgoing AS-path filters to the BGP neighbor.

D) Force the updates to go through the new filter.

Practice: Monitoring AS-Path Filters

Q1) How can you test your regular expression?

A) **show ip bgp access-list** command

B) **show ip bgp filter** command

C) **show ip bgp regexp** command

D) **show ip bgp summary** command

Q2) How can you test your AS-path filter before using it?

A) **show ip bgp access-list** command

B) **show ip bgp filter** command

C) **show ip bgp regexp** command

D) **show ip bgp summary** command

 Copyright © 2003, Cisco Systems, Inc.

Prefix-List Filters

Practice: Requirements for Prefix-Based Filters

Q1) What are two reasons that a multihomed customer needs prefix-lists? (Choose two.)

A) to ensure that only valid IP prefixes are announced to the ISPs

B) to set a limit on the number of prefixes that it can accept from the ISPs

C) to prevent the customer from receiving its own IP prefixes from the ISP

D) to verify that the customer has received full Internet route tables

Practice: Prefix-Lists vs. IP Access-Lists

Q1) What are three benefits of prefix-lists when compared to IP access-lists? (Choose three.)

A) Prefix-lists have a tree structure so the "permit" or "deny" result can be reached faster.

B) Prefix-lists can indicate that the subnet mask must be within a specified range.

C) If none of the lines in a prefix-list match, the result is "implicitly deny."

D) Individual entries in prefix-lists are easier to insert or delete.

Practice: Configuring Prefix-Lists

Q1) When you are defining prefix-lists, what are two reasons to use sequence numbers? (Choose two.)

A) to reference the associated ACL for the prefix-list entry

B) to provide a means of linking an AS-path filter-list to the prefix-list

C) to provide an execution order for prefix-list entries

D) to provide a means of inserting or deleting list entries

Copyright © 2003, Cisco Systems, Inc.

Q2) When you are defining a prefix-list using the *ge-value* or *le-value*, what is true regarding the *len* parameter?

A) *ge-value* is greater than or equal to *le-value*

B) *le-value* is less than or equal to *len*

C) *len* is less than *ge-value* is less than or equal to *le-value* is less than or equal to 32

D) 8 is less than or equal to *len* is less than or equal to *ge-value* is less than or equal to 32

Practice: BGP Filters Implementation

Q1) How are prefix-lists and the AS-path filters combined?

A) They cannot be combined, because the router can use either prefix-lists or AS-path filters but not both simultaneously.

B) The prefix-list and AS-path lists use the same filter-list.

C) The router checks both lists for the first match and then stops processing.

D) All updates are matched against the prefix-list and the AS-path filter before being sent or accepted.

Practice: Implementing Prefix-Lists in the BGP Process

Q1) How can you apply the same prefix-list to multiple BGP neighbors on a router?

A) by configuring a **neighbor prefix-list** statement for each BGP peer

B) by configuring a **neighbor distribute-list** statement for each neighbor

C) by using the BGP *peer-group* option with the **neighbor** statement

D) by configuring the prefix-list as a global filter under the BGP routing process

 Copyright © 2003, Cisco Systems, Inc.

Practice: Modifying Prefix-Lists

Q1) If sequence number 2 is inserted in the first line of a prefix-list, but no other sequence numbers are entered on subsequent lines, what will the sequence numbers be for the next three statements?

A) This will generate an error, because all remaining sequence numbers must be manually entered.

B) 3, 4, 5

C) 7, 12, 17

D) 5, 10, 15

Practice: Monitoring Prefix-Lists

Q1) How can you use the **show ip prefix-list** command to display the prefix-list entry matching a specific prefix and length?

A) This is not a feature of the **show ip prefix-list** command.

B) By specifying the **detail** keyword.

C) With the longer keyword to display all matches except those with more specific entries.

D) By specifying the **first-match** keyword.

Outbound Route Filtering

Practice: Outbound Route Filter

Q1) What are two benefits of using the outbound route filtering feature? (Choose two.)

A) It minimizes the amount of BGP routing traffic transferred over the directly connected link.

B) It defines inbound filters used on remote peer routers.

C) It reduces the amount of CPU consumed by a router.

D) You can use ORF to implement nondisruptive changes in BGP policy.

Practice: Outbound Route Filter Message

Q1) What ORF type is currently supported on Cisco routers?

A) 1

B) 2

C) 3

D) 128

Q2) What are three key components of an ORF type 1 message? (Choose three.)

A) action clause

B) match clause

C) set clause

D) when clause

 Copyright © 2003, Cisco Systems, Inc.

Practice: Configuring Outbound Route Filtering

Q1) How is capability negotiation for prefix-list–based ORF enabled?

A) by configuring a prefix-list filter on the router and enabling BGP soft route refresh

B) by issuing a **clear ip bgp neighbor** command on the router

C) by configuring both neighbors with complementing ORF capabilities using the **neighbor capability orf prefix-list** command

D) by applying a prefix-list to the BGP session with the **neighbor prefix-list** command

Practice: Using Outbound Route Filtering

Q1) What is the recommended method of triggering an ORF route refresh?

A) Reload the remote peer router.

B) Remove and replace the BGP neighbor configuration.

C) Execute the **clear ip bgp *** command on the router.

D) Execute the **clear ip bgp** *neighbor* **in [prefix-filter]** command.

Practice: Monitoring Outbound Route Filtering

Q1) What command can you use to verify that a router supports the exchange of ORF capabilities?

A) **show ip bgp summary**

B) **show cdp neighbors**

C) **show ip bgp neighbors**

D) **show ip bgp orf capabilities**

Copyright © 2003, Cisco Systems, Inc.

Route-Maps as BGP Filters

Practice: Why Use Route-Maps as BGP Filters?

Q1) What are two reasons for using route-maps on BGP neighbors? (Choose two.)

A) when you want to change the attributes of BGP updates

B) when you want to filter BGP updates based on a complex set of conditions

C) when you want to implement a simple routing policy for incoming and outgoing BGP updates

D) when you want to use both a prefix-list and an AS path to filter BGP updates

Practice: Route-Maps Overview

Q1) Which BGP attributes can you match with a route-map?

A) AS-path; community; next-hop; origin; route originator

B) AS-path; community; next-hop; weight; route originator

C) AS-path; local preference; MED; next-hop; route

D) Community; MED; next-hop; origin; prefix

Q2) Which BGP attributes can you set with a route-map?

A) AS-path; community; next-hop; origin; route originator

B) AS-path; local preference; MED; next-hop; route originator

C) Community; local-preference; MED; next-hop; origin

D) Community; local-preference; MED; next-hop; route originator

 Copyright © 2003, Cisco Systems, Inc.

Practice: Prefix-List Use in Route-Maps

Q1) What are two ways in which route-maps are combined with prefix-lists and AS-path filters? (Choose two.)

A) You can apply route-maps, AS-path filters, and prefix-lists to incoming or outgoing routes in any combination.

B) You can use route-maps with AS-path filters but not with prefix-lists.

C) You can use prefix-lists and AS-path filters within the **match** statements of a route-map.

D) You cannot combine route-maps with prefix-lists or AS-path access-lists.

Practice: BGP Filters

Q1) What are three situations in which you can apply route-maps on a router to enable BGP policy enforcement? (Choose three.)

A) to apply policy to inbound BGP updates from a neighbor

B) to apply policy to outbound BGP updates sent to a neighbor

C) to apply policy to locally originated routes

D) to filter routes distributed from an IGP into BGP

Practice: Using Route-Maps as BGP Filters

Q1) How do you use a route-map with BGP?

A) Route-maps are applied globally to the BGP routing process.

B) Route-maps are applied to BGP neighbors in the inbound or outbound direction.

C) Route-maps are applied to the interface connecting EBGP peers.

D) Route-maps are automatically appended to configured BGP neighbors.

Copyright © 2003, Cisco Systems, Inc.

Practice: Monitoring Route-Maps

Q1) How does a **set** statement in an outgoing route-map affect the BGP table?

A) A copy of the modified BGP entry is stored in the BGP table but is not used.

B) The result of applying the filters to the IP routing table is applied to the BGP table.

C) The changes are reflected in the BGP table.

D) It does not affect the local BGP table in any way.

 Copyright © 2003, Cisco Systems, Inc.

Implementing Changes in BGP Policy

Practice: Traditional Filtering Limitations

Q1) Why is clearing a BGP session a disruptive change in routing policy?

A) Clearing a BGP session takes a long time and can disrupt packet forwarding.

B) You cannot recover information sent while the BGP session is being cleared.

C) You cannot automatically re-establish sessions that are torn down during the clearing operation.

D) You cannot selectively tear down BGP sessions; you must clear sessions with all neighbors.

Practice: BGP Soft Reconfiguration

Q1) What is the impact of inbound soft reconfiguration?

A) It clears the session after you reconfigure the new routing policy.

B) It creates a copy of all routes received from a neighbor after the filters are applied.

C) It requires extra memory to hold a copy of all routes received from the neighbor.

D) It resets the table version number of the neighbor to 0.

Q2) What are two impacts of outbound soft reconfiguration? (Choose two.)

A) No extra CPU resources will be consumed.

B) No extra memory will be consumed.

C) Extra memory is required, because a copy of the BGP table is maintained.

D) Extra CPU resources are consumed only while the **clear ip bgp** command is being issued.

Copyright © 2003, Cisco Systems, Inc.

Practice: Cisco IOS Commands for Soft Reconfiguration

Q1) What two steps must you complete to use inbound soft configuration functionality? (Choose two.)

A) Clear the BGP session inbound in the local router.

B) Clear the BGP session outbound on the remote router.

C) Configure the local neighbor with **soft-reconfiguration in**.

D) Configure the remote neighbor with **soft-reconfiguration out**.

Q2) If the **in** or **out** option is not specified when a network administrator is using soft reconfiguration, what does the router do?

A) The router will initiate a soft reconfiguration inbound.

B) The router will initiate a soft reconfiguration outbound.

C) The router will initiate a soft reconfiguration both inbound and outbound.

D) The router will not initiate a soft reconfiguration until either the inbound or outbound option is specified.

Practice: Monitoring Soft Reconfiguration

Q1) How do you check the stored BGP table when you are using inbound soft reconfiguration?

A) With the **show ip bgp neighbor** *ip-address* **advertised** command.

B) With the **show ip bgp summary route advertised** command.

C) With the **show ip bgp neighbor** *ip-address* **received** command.

D) There is no means to display the stored copy of the BGP table before soft reconfiguration.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Route Refresh

Q1) What are two situations when would you prefer inbound soft reconfiguration to route refresh? (Choose two.)

A) when there is insufficient memory to hold a copy of the BGP table of the neighbor

B) when a route refresh fails

C) when you wish to troubleshoot filters and use the **show ip bgp neighbor** command with the **received-routes** option

D) when the neighboring router does not support the route refresh capability

Practice: Using Route Refresh

Q1) What command do you use to trigger route refresh to a specific BGP neighbor?

A) **clear ip bgp** *ip-address* **in**

B) **show ip bgp neighbor**

C) **clear ip bgp ***

D) **clear ip bgp** *ip-address* **soft out**

Practice: Monitoring Route Refresh

Q1) How do you determine whether a BGP neighbor supports route refresh?

A) A flag in the BGP table indicates the presence of route refresh capability.

B) The **show ip bgp neighbor** command indicates if the option is supported.

C) Initiate **debug ip bgp negotiation** to see if the router has completed a route refresh capabilities exchange.

D) Execute the **clear ip bgp *** command. Command-line BGP status messages will indicate route refresh support capabilities.

 Copyright © 2003, Cisco Systems, Inc.

Module 3: Route Selection Using Attributes

BGP Path Attributes

Practice: BGP Path Attributes

Q1) What do network administrators use in BGP to define the metrics used for best route selection?

A) link states

B) path attributes

C) distance vectors

D) cost

Practice: Well-Known BGP Attributes

Q1) Which three BGP path attributes must be carried with each update? (Choose three.)

A) origin

B) AS-path

C) local preference

D) next-hop

Q2) Which two attributes are discretionary well-known BGP attributes? (Choose two.)

A) MED

B) local preference

C) origin

D) atomic aggregate

Copyright © 2003, Cisco Systems, Inc.

Practice: Optional BGP Attributes

Q1) How are recognized transitive optional attributes propagated between BGP neighbors?

A) Optional attributes are converted to transitive well-known attributes.

B) With the partial bit set.

C) Based on their meaning.

D) Optional attributes are not propagated to neighbors.

Q2) How are nonrecognized transitive optional attributes propagated between BGP neighbors?

A) Optional attributes are converted to transitive well-known attributes.

B) With the partial bit set.

C) Based on their meaning.

D) Optional attributes are not propagated to neighbors.

Q3) Which is a nontransitive optional BGP path attribute?

A) local preference

B) weight

C) MED

D) community

Practice: AS-Path Attribute

Q1) How do BGP routers detect routing loops?

A) BGP routers check for the longest AS path in all routing updates.

B) BGP routers ignore incoming routes with an AS path containing their own AS number.

C) Multiple paths with the same origin AS are considered routing loops.

D) Multiple paths with the same destination AS are considered routing loops.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Next-Hop Attribute

Q1) When is the next-hop attribute different from the IP address of the sending router?

A) In situations where static routing is used to reach nondirectly connected IBGP peers.

B) In cases where the BGP next hop is in a different subnet than the receiving router.

C) The next-hop attribute is always different, because it points to the exit gateway.

D) If the current BGP next hop is in the same IP subnet as the receiving router.

Q2) How can improperly designed, partially meshed networks break BGP connectivity?

A) A direct connection may not be available to the same subnet next hop.

B) A design using separate IP networks and subinterfaces can cause BGP to set the incorrect next hop.

C) The BGP next hop is always set to the sending router, causing it to act as the hub of the meshed network.

D) Partially meshed networks automatically disable BGP next-hop processing and should be avoided.

Copyright © 2003, Cisco Systems, Inc.

Influencing BGP Route Selection with Weights

Practice: BGP Route Selection Criteria

Q1) What is the difference between local preference and weight?

A) Local preference has a higher priority in BGP path selection.

B) Local preference is used AS-wide while weight is local to a single router.

C) Local preference is local only to a specific BGP-speaking router.

D) Local preference is used to influence incoming path selection.

Practice: Influencing BGP Route Selection

Q1) What three methods can you use to set the BGP weight attribute? (Choose three.)

A) route-map

B) AS-path filter-list

C) access control list

D) default weight assigned to a specific neighbor

Practice: Configuring Per-Neighbor Weights

Q1) What is the default weight for routes received from a BGP neighbor?

A) 0

B) 100

C) 32768

D) depends on the Cisco IOS release

 Copyright © 2003, Cisco Systems, Inc.

Q2) What default weight is applied to locally sourced BGP routes?

A) 0

B) 100

C) 32768

D) depends on the Cisco IOS release

Q3) When are the weights configured on a neighbor enforced?

A) Before the new weights can take effect, the BGP process on the router must be removed and reconfigured.

B) The router must first be rebooted for the new weights to take effect.

C) The new weights will be applied after the BGP update interval of 30 minutes expires.

D) The new weight configuration is applied to all routes received following the configuration change.

Practice: Changing Weights with AS-Path Filters

Q1) How could you implement a primary/backup ISP routing policy using weights?

A) Assign higher weights to all routes received from the backup ISP.

B) Assign lower weights to all routes received from the backup ISP.

C) Assign higher weights to all routes received from the primary ISP.

D) Assign lower weights to all routes received from the primary ISP.

Q2) What is the difference between a **filter-list in** configuration command and a **filter-list weight** configuration command?

A) The filter-list without the **weight** keyword can manipulate AS paths or the weight attribute, but the **weight** keyword is limited to the weight attribute.

B) **Filter-list weight** will not filter out any route but will assign the weight value to those routes permitted by the filter-list.

C) Using **filter-list weight** forces the router to use process switching.

D) **Filter-list weight** can match an AS path with a single AS number entry only.

Practice: Changing Weights with Route-Maps

Q1) When you are using route-maps to modify weights, what happens by default to a route that does not match any of the route-map statements?

A) The route is accepted with the weight attribute unmodified.

B) The route is discarded.

C) The route will be inserted into the BGP table but not the IP routing table.

D) An error will be displayed on the router console and in router debugs.

Practice: Monitoring BGP Route Selection and Weights

Q1) What three pieces of information can you obtain from the output of the **show ip bgp** command? (Choose three.)

A) the best route to a destination

B) the weight attribute

C) the AS path of the route

D) the administrative distance of the route

Practice: BGP Route Selection and Filtering Tools Summary

Q1) Which method of influencing route selection with weights is the last to be applied on an incoming interface?

A) prefix-list

B) route-map

C) filter-list weight

D) default weight

 Copyright © 2003, Cisco Systems, Inc.

BGP Local Preference

Practice: Consistent Route Selection Within the AS

Q1) What is a key difference between the local preference and weight attributes?

A) Local preference is local to the route on which it is configured.

B) Local preference is local to the AS within which it has been configured.

C) Local preference is local to the BGP administrative domain.

D) Local preference is global to a BGP domain.

Q2) What is an appropriate BGP implementation for the weight attribute?

A) all large scale BGP implementations requiring AS-wide path selection policies

B) simple routing policies in smaller networks

C) BGP implementations connecting a single-homed customer to provider

D) domains requiring global path selection policies

Practice: BGP Local Preference

Q1) If you configure both local preference and weight, which has the highest priority?

A) Local preference always has a higher priority.

B) Weight always has a higher priority.

C) If routes are learned from different neighbors, the local preference will determine path selection.

D) If two routes have the same local preference, weight will determine the selected path.

Copyright © 2003, Cisco Systems, Inc.

Practice: Configuring Default Local Preference

Q1) What is the default value of local preference?

A) 0

B) 100

C) 255

D) 32768

Practice: Configuring Local Preference with Route-Maps

Q1) What effect does a route-map have on the local preference setting of outgoing EBGP updates?

A) The local preference can be set for outgoing EBGP updates to notify the neighboring AS about desired path selection.

B) Route-maps cannot be used to apply changes to local preference in the outbound direction.

C) You can configure local preference only on IBGP neighbors.

D) Applying a route-map to outgoing updates on EBGP sessions does not affect local preference in the neighboring AS.

Practice: Monitoring Local Preference

Q1) What command should you use to display locally applied local preference settings?

A) **show bgp preference detail**

B) **show ip bgp**

C) **show ip bgp detail**

D) **show ip bgp** *prefix*

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Prepending

Practice: Return Path Selection in a Multihomed AS

Q1) What are two important reasons for influencing return path selection in a multihomed AS? (Choose two.)

A) to prevent asymmetrical traffic flow of traffic returning to the AS

B) to prevent network replay and denial-of-service attacks

C) to prevent suboptimal performance of return traffic that prefers to use lower-bandwidth links

D) to eliminate the need to influence outgoing path selection using BGP attributes

Q2) In a multihomed scenario, why does BGP not always use the highest-bandwidth link available?

A) because most administrators fail to configure interface bandwidths using the **bandwidth** command

B) because, if BGP is not configured to properly redistribute into the IGP, bandwidth metrics are lost from their attached routes

C) because many administrators fail to properly set the metric on each route as a reflection of the link bandwidth

D) because, if not influenced with attributes, the shortest AS path decides which path is used by BGP traffic

Practice: AS-Path Prepending

Q1) What are two functions of AS-path prepending? (Choose two.)

A) to mask the actual origination point of a route

B) to influence the path that a route takes out of its originating AS

C) to influence the path that a route takes into its originating AS

D) to distribute return traffic load between multihomed routers

Q2) What will happen if you prepend an AS number other than that of the originating AS to the AS path?

A) The upstream AS will automatically reject the route.

B) The route will traverse the Internet and be discarded by the downstream provider.

C) If the route is sent to the AS matching the prepended string, loop prevention mechanisms will silently ignore the update.

D) The AS-path prepending function will allow only the originating AS to be prepended to the AS path.

Practice: AS-Path Prepending Design Considerations

Q1) What are two design principles for deciding how many AS number strings to prepend to an AS path? (Choose two.)

A) The AS-path length should be equal to one plus the number of AS-path hops to the destination, if load balancing is desired.

B) The number of AS numbers to append should equal the number of multihomed connections plus one.

C) If load distribution is desired, append as few AS numbers as possible to achieve the desired balancing goal.

D) If a backup link is desired, create a large AS path to ensure that the primary link is always used.

Practice: Configuring AS-Path Prepending

Q1) When you are configuring AS-path prepending with route-maps, when do the changes to the AS path take effect for already announced routes?

A) The changes are applied immediately.

B) The changes will take effect upon clearing the BGP neighbor session of the configured neighbor.

C) The changes will take effect only after reloading the router.

D) Changes are applied only after the upstream AS accepts the AS path.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Monitoring AS-Path Prepending

Q1) Where is the best place to monitor the AS-path prepending operation?

A) on the sending router using the **debug ip bgp updates** command

B) on the sending router using the **show route-map** command

C) on the receiving router using the **show ip bgp regexp** command

D) on either the sending or receiving router

Practice: AS-Path Filtering Concerns with AS-Path Prepending

Q1) How may the AS of a service provider have to change its configurations when customers manipulate their outgoing AS paths?

A) Upstream providers have to disable all AS-path filtering.

B) The upstream provider must modify any filters based on the AS-path attribute to allow the prepended path.

C) The provider has to create an AS-path filter specific to each customer.

D) Customer-manipulated AS paths typically require no changes in the provider network.

Copyright © 2003, Cisco Systems, Inc.

BGP Multi-Exit Discriminator

Practice: Selecting the Proper Return Path

Q1) What is the typical application of the MED attribute?

A) to influence path selection out of an originating AS

B) to provide a strong metric to select the best path when multiple routes exist

C) to have a BGP attribute traversing many autonomous systems while influencing path selection

D) to influence the return path of traffic back into an AS

Practice: MED Propagation in a BGP Network

Q1) What happens to the MED attribute when it is sent to EBGP peers?

A) The value of MED is ignored and reset to 0.

B) All MED values are carried into an AS but are removed by the neighbor AS.

C) The MED is carried into and sent to all EBGP neighbors.

D) MED values originating in the AS are carried into the neighboring AS but do not leave it.

Practice: Changing Default MED

Q1) What is the default value of MED?

A) 0

B) 100

C) 32768

D) MED is unused unless explicitly defined.

 Copyright © 2003, Cisco Systems, Inc.

Q2) What is the default value of MED when routes are redistributed from another routing protocol?

A) MED is set to 0 for all redistributed routes.

B) MED is set to 100 for all redistributed routes.

C) There is no MED attribute attached to a redistributed route.

D) MED is not attached to redistributed routes except for those matching the IP routing table.

Practice: Changing MED with Route-Maps

Q1) What are three BGP attributes that are compared before MED? (Choose three.)

A) largest weight

B) originated routes

C) AS-path length

D) lowest IP address

Practice: Advanced MED Configuration

Q1) What effect does the **bgp deterministic-med** command have on BGP path selection?

A) It forces a MED attribute comparison between MED attributes from different neighbor autonomous systems.

B) It prohibits MED attribute comparison if routes are from a different AS.

C) It ensures that an accurate MED comparison is made across all routes received from the same AS.

D) It allows the MED to determine BGP path selection above other BGP attributes.

Practice: Monitoring MED

Q1) What two statements are true regarding MED in **show ip bgp** commands? (Choose two.)

A) MED is always displayed.

B) MED is displayed only for those routes having a MED attribute.

C) MED is listed as a metric value.

D) MED is listed as a MED value.

Practice: Troubleshooting MED

Q1) If you configure inbound soft reconfiguration with a route-map and issue the **show ip bgp** *prefix* command, which value of the MED attribute is displayed?

A) Only the original route (no MED) is displayed.

B) Both the original route and the modified route are displayed.

C) Only the modified route is displayed.

D) The MED attribute is not displayed with the **show ip bgp** *prefix* command.

 Copyright © 2003, Cisco Systems, Inc.

BGP Communities

Practice: Selecting the Proper Return Path

Q1) What are two reasons why it is not feasible to use MED to influence return path selection when multiple autonomous systems are involved? (Choose two.)

A) because the MED attribute is designed to influence outbound path selection only

B) because the AS-path attribute would be used for path selection regardless of any configured MED value

C) because the weight attribute will always be used, given that it is first in the BGP route selection process

D) because MED cannot be propagated across several autonomous systems

Practice: BGP Communities Overview

Q1) What is the purpose of the community attribute?

A) to filter incoming BGP route updates

B) to update the BGP table with incoming BGP routes

C) to facilitate selection of the optimum AS exit path

D) to group destinations by tagging BGP updates

Practice: Using Communities

Q1) Does the community attribute have any influence on BGP path selection?

A) No, communities are simply tags that are applied to BGP routes.

B) No, communities are nontransitive attributes.

C) Yes, BGP paths are selected based on the value in the community tag.

D) Yes, the community attribute is part of the BGP route selection process.

Copyright © 2003, Cisco Systems, Inc.

Practice: Configuring BGP Communities

Q1) What Cisco IOS feature must you use to set the BGP community attribute?

A) distribute-list

B) filter-list

C) route-map

D) access control list

Practice: Configuring Route Tagging with BGP Communities

Q1) How many community tags can be attached to a single BGP route?

A) 1

B) 32

C) 255

D) depends on the number configured with the **ip bgp community** command

Practice: Configuring Community Propagation

Q1) What must you configure on a Cisco router to begin the propagation of the community attribute?

A) a BGP peer group

B) **send-community** on the **neighbor** statement.

C) a distribute-list under the router BGP process

D) an outgoing route-map on the **neighbor** statement

 Copyright © 2003, Cisco Systems, Inc.

Practice: Defining BGP Community-Lists

Q1) What match criteria are specified in a standard BGP community-list?

A) destination IP addresses

B) regular expressions

C) community attribute values

D) AS numbers

Q2) What regular expression should you use with an extended BGP community-list to match any community value?

A) internet

B) any

C) .*

D) permit all

Practice: Matching BGP Communities in Route-Maps

Q1) What is the result of tagging a route with the no-export community?

A) The route will not be advertised within the local AS.

B) The upstream AS will not be allowed to export the route.

C) The route cannot be exported to another routing protocol.

D) The router will not propagate the route to any external neighbors except to intraconfederation external neighbors.

Copyright © 2003, Cisco Systems, Inc.

Practice: Monitoring Communities

Q1) What command should you use to display the community attribute?

A) **show ip bgp summary**

B) **show community**

C) **show ip bgp** *prefix*

D) **show ip route**

 Copyright © 2003, Cisco Systems, Inc.

Module 4: Customer-to-Provider Connectivity with BGP

Customer-to-Provider Connectivity Requirements

Practice: Customer Connectivity Types

Q1) Which method of Internet connectivity gives the highest level of redundancy?

A) a single permanent connection to one ISP

B) multiple permanent connections to one ISP where one of the lines is primary and the other line is used for backup only

C) multiple permanent connections to one ISP used for load sharing of traffic

D) multiple permanent connections to more than one ISP

Practice: Redundancy in Customer Connections

Q1) Multiple permanent connections to a single ISP do not provide what level of redundancy?

A) redundancy on link failure

B) redundancy on ISP failure.

C) redundancy on equipment failure

D) redundancy on routing protocol failure

Q2) What is a drawback when a customer is installing multiple permanent connections to multiple ISPs?

A) There is no redundancy on ISP failure.

B) Good load sharing is impossible to achieve.

C) The customer can use only Frame Relay PVCs.

D) Equipment failure may cause a complete network outage.

Copyright © 2003, Cisco Systems, Inc.

Practice: Customer-to-Provider Routing Schemes

Q1) Which form of routing provides the best redundancy?

A) static routing

B) content routing

C) dynamic routing

D) embryonic routing

Practice: Customer Routing

Q1) What will happen if a link failure is not detected where multiple permanent connections between a single router on the customer side and a single router on the ISP side are configured with static routes?

A) Fast failover will occur.

B) The traffic will be black-holed.

C) Customers will get a message that the Internet is down.

D) Nothing happens. All routing continues as usual.

Q2) Multiple permanent connections to more than one ISP always require the use of what?

A) dynamic routing

B) mobile routing

C) static routing

D) secure routing

Practice: Addressing Requirements

Q1) When a customer decides to use private addresses within its network, what must it do to connect to its ISP?

A) No action is required.

B) The customer must summarize its address space.

C) The customer must configure static addressing.

D) The customer must use NAT.

 Copyright © 2003, Cisco Systems, Inc.

Practice: AS Number Allocation

Q1) What number ranges correctly indicate private AS numbers?

A) 1024 – 2048

B) 32768 – 64511

C) 64512 – 65535

D) 65536 – 131072

Implementing Customer Connectivity Using Static Routing

Practice: Why Use Static Routing?

Q1) Which of these two situations identify when to use static routing between a customer and a service provider in a BGP environment? (Choose two.)

A) customers with a single connection to the Internet

B) customers with multiple connections to multiple service providers

C) customers with multiple connections to the same service provider

D) customers with a single connection to multiple service providers

Practice: Characteristics of Static Routing

Q1) What could potentially cause poor performance of the service provider IGP routing?

A) when static routing is used

B) redistribution of customer routes into the IGP

C) if a customer uses PA addresses

D) if the edge router of the customer announces itself as a default gateway

Practice: Designing Static Route Propagation in a Service Provider Network

Q1) What must you identify before you configure a static route for a specific customer?

A) the static routes

B) the corresponding tag value

C) the corresponding community value

D) the combination of the services provided to this customer

 Copyright © 2003, Cisco Systems, Inc.

Practice: Static Route Propagation Case Study Parameters

Q1) How many tag value(s) must you assign to each combination of services?

A) 1

B) 2

C) 16

D) 32

Q2) What do you use at the edge router of a service provider to redistribute static routes into BGP?

A) **no-export**

B) **send-community**

C) **route-map**

D) **route-community**

Practice: BGP Backup with Static Routes

Q1) What is configured on backup routers and redistributed into the customer IGP and provider BGP after a primary link fails?

A) weighted routes

B) floating static routes

C) floating dynamic routes

D) dynamic forwarding table

Q2) When you are configuring a backup router with BGP static routes, how can you ensure that as long as the backup router is receiving the default route from the IGP that the BGP static route is not used?

A) Assign the static route a low AD value.

B) Assign the static route a high AD value.

C) Assign the static route an AD equal to that of the IGP.

D) Nothing, the route selection process is automatic.

Copyright © 2003, Cisco Systems, Inc.

Practice: Floating Static Routes with BGP

Q1) Which two of the following statements indicate limitations of floating static routes within BGP? (Choose two.)

A) Weight values cannot be modified.

B) Floating routes do not work correctly.

C) AD cannot be matched with a route-map.

D) There is no way to define additional tags for static routes.

Practice: Load Sharing with Static Routes

Q1) When you are performing load sharing of outgoing traffic with static routes, what is the effect of colocating the edge routers?

A) greater throughput

B) faster convergence

C) higher availability

D) balanced load sharing

Q2) When you are using static routes, with what two routing tricks can you optimize return traffic load sharing? (Choose two.)

A) Each provider edge router advertises only part of the customer address space into the provider backbone.

B) Each provider edge router advertises the entire address space of the customer into the provider backbone.

C) Each provider edge router advertises the entire address space of the customer for backup purposes.

D) Each provider edge router advertises only part of the customer address space for backup purposes.

 Copyright © 2003, Cisco Systems, Inc.

Connecting a Multihomed Customer to a Single Service Provider

Practice: Configuring BGP on Multihomed Customer Routers

Q1) Why can you not use static routing in all cases of redundant links between a customer and a single ISP?

A) You cannot use static routing when multihoming to a single provider.

B) You cannot use static routing when load balancing is a design requirement.

C) You cannot use static routing when the ISP conditionally advertises customer routes.

D) You cannot use static routing in scenarios where the failures cannot be detected by Layer 2 protocols.

Practice: Conditional Advertising in Multihomed Customer Networks

Q1) Why should you not configure conditional route advertising using a route to null 0?

A) Because the route to null 0 will cause all traffic to de discarded.

B) A route to null 0 will condition BGP to advertise routes only when the interface with a matching assigned network number is in the up state.

C) A route to null 0 will condition BGP such that the network will always be advertised regardless of its state.

D) Using null 0 as a conditional advertisement will advertise the network only if a matching route exists in the IP routing table.

Practice: Configuring BGP on Service Provider Routers

Q1) Functionally, what three requirements must you take into consideration when configuring in the service provider network to support a multihomed customer? (Choose three.)

A) The provider should filter incoming BGP updates with a prefix-list to verify customer-announced assigned space.

B) The provider should filter incoming BGP updates with an AS-path filter-list to verify that the customer uses its own AS number only.

C) The provider should advertise a default route to the customer through BGP.

D) The provider should announce nonsummarized prefixes for the customer networks to the Internet.

Practice: Removing Private AS Numbers

Q1) AS 64525 is connected to AS 229, which in turn is connected to AS 1126. How will configuring private AS removal on AS 229 affect routers in AS 1126?

A) There will be no effect on the routers in AS 1126.

B) The routers in AS 1126 will see all Internet routes as originating in AS 64525.

C) The routers will see networks originating in AS 64525 as originating in AS 229.

D) AS 1126 will receive all of its routes with an AS-path length of two.

Practice: Backup Solutions with BGP

Q1) Which attribute can you use to select the primary/backup link for outgoing traffic?

A) weight

B) local preference

C) AS-path

D) MED

 Copyright © 2003, Cisco Systems, Inc.

Q2) Which attribute can you use to select the primary/backup link for incoming traffic?

A) weight

B) local preference

C) AS-path

D) MED

Practice: Load Sharing with the Multihomed Customer

Q1) What three options can you use to enable load sharing on parallel links connected to one router? (Choose three.)

A) Split the customer address space into two parts and advertise a portion on each link.

B) Use AS-path prepending on the outgoing routes of the backup path.

C) Use the **ebgp-multihop** option between loopback interfaces of the multihomed routers.

D) Enable BGP multipath support on the multihomed routers.

Practice: Load Sharing with BGP Multipath

Q1) By default, BGP can perform load balancing over how many parallel links?

A) one

B) four

C) six

D) eight

Practice: Load Sharing with EBGP Multihop

Q1) What three criteria must be met before two routers with parallel links between them can perform load balancing using **ebgp-multihop**? (Choose three.)

A) The **neighbor ebgp-multihop** command must be configured on each router.

B) BGP maximum paths must be set to the number of links between the routers.

C) The routers must have a static route or an IGP containing reachability information for the configured loopback addresses.

D) The **neighbor update-source** option must be configured on both routers.

 Copyright © 2003, Cisco Systems, Inc.

Connecting a Multihomed Customer to Multiple Service Providers

Practice: Configuring BGP for Multihomed Customers

Q1) Why should you apply outbound filters to a multihomed customer network?

A) to provide for maximum security at the customer site

B) to guarantee that return traffic has proper load balancing

C) to prevent the customer network from becoming a transit AS

D) to ensure that outbound traffic has proper load balancing from the customer site

Practice: Multihomed Customer Address Space Selection

Q1) Why is PI address space recommended for customers connected to multiple providers?

A) PI addressing facilitates easier implementation of backup link policies.

B) PI addressing facilitates easier implementation of load-sharing configurations.

C) PI addressing removes any limitations around advertising the address space.

D) PI addressing enables the use of automatic summarization at upstream ISPs.

Practice: Multihomed Customer AS Number Selection

Q1) What are three AS number implementation options available to customers connecting to multiple providers? (Choose three.)

A) The customer can obtain a registered, public AS number that is advertised to all upstream providers.

B) The customer can use a single private AS number as long as all upstream providers agree to support the same AS number.

C) The customer can use two different private AS numbers by translating one of the private addresses at the customer edge.

D) The customer can use two different PA AS numbers by configuring EBGP internally at the customer site.

Copyright © 2003, Cisco Systems, Inc.

Practice: AS Number Translation

Q1) What are two issues that arise when you are using AS number translation? (Choose two.)

A) The upstream provider must not filter routes based on a single AS.

B) AS-path prepending is not supported when you are using AS number translation.

C) AS number translation causes two AS numbers to be prepended to the AS path.

D) You can use only private AS numbers with AS number translation.

Practice: Primary/Backup Link Selection

Q1) In what two ways is the primary/backup design different from the one used for multihomed customers connected to a single ISP? (Choose two.)

A) You cannot accomplish incoming route selection using the MED attribute.

B) You can accomplish outgoing route selection only by using weights.

C) The customer should use local preference to direct traffic to the correct outgoing link.

D) BGP communities and AS-path prepending are used to influence incoming route selection.

Practice: BGP Incoming Link Selection

Q1) What are three benefits of using AS-path prepending instead of BGP communities when you are influencing incoming path selection? (Choose three.)

A) The customer is not dependent upon the service provider for configuration changes and maintenance.

B) Using BGP communities to set local preference may not work in all network scenarios.

C) AS-path prepending requires only that the backup provider support multiple instances of the same AS number in the AS path to function correctly.

D) Using BGP communities to set local preference cannot be scaled in a service provider network.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Load Sharing with Multiple Providers

Q1) In what two ways is the load-sharing design different from the one used for multihomed customers connected to a single ISP? (Choose two.)

A) There is no difference in the design for outgoing traffic.

B) There is no difference in the design for incoming traffic.

C) You can use AS-path prepending.

D) The weight, MED and local preference attributes all have to be altered.

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

Module 5: BGP Transit Autonomous Systems

Transit Autonomous System Functions

Practice: Transit Autonomous System Tasks

Q1) What are the two main functions of a transit AS? (Choose two.)

A) to propagate routes between remote autonomous systems

B) to filter noncustomer routes from transiting the AS

C) to route packets between remote networks

D) to connect customer networks to Internet service providers

Practice: External Route Propagation

Q1) How are external routes exchanged between autonomous systems?

A) through EBGP

B) through IBGP

C) through route redistribution into an IGP

D) through local advertisement of routes at the edge router

Practice: Internal Route Propagation

Q1) How are BGP routes propagated across an AS?

A) through EBGP

B) through IBGP

C) through route redistribution into an IGP

D) through local advertisement of routes at the edge router

Copyright © 2003, Cisco Systems, Inc.

Practice: Packet Forwarding in an Autonomous System

Q1) Why do you have to run BGP on all core routers in a transit AS?

A) to eliminate the possibility of routing loops

B) to optimize the routing across the AS

C) to be able to forward packets to all external destinations

D) to ensure that only one exit point exists for the transit backbone

Practice: Core Router IBGP Requirements in a Transit Autonomous System

Q1) Why is the redistribution of BGP routes into IGP not advisable?

A) Redistributing full Internet routing into any IGP would result in the loss of the BGP attributes needed to ensure optimal routing within an AS.

B) Redistributing full Internet routing is not possible, because BGP policies cannot be enforced by IGPs.

C) IGPs are not capable of handling the number of routes currently present on the Internet.

D) The increased convergence times of IGPs as compared to BGP would cause too many flaps, rendering BGP unstable for Internet use.

 Copyright © 2003, Cisco Systems, Inc.

IBGP and EBGP Interaction in a Transit Autonomous System

Practice: AS-Path Processing in IBGP

Q1) How is the AS-path attribute modified across a transit AS?

A) The AS path of each router is prepended to the AS-path attribute.

B) The AS path is not changed when the BGP prefix is propagated across IBGP sessions.

C) The AS number of the router sending the information is appended to the AS path.

D) The value of the AS path is reset to the AS path of the transit backbone.

Practice: BGP Split Horizon

Q1) What is the purpose of the BGP split-horizon rule?

A) to propagate IBGP updates to other IBGP peers

B) to prevent BGP routing loops within an AS

C) to send routing information to an EBGP neighbor

D) to prevent BGP routes from being sent back to the advertising router

Q2) What is the impact of BGP split horizon on an AS?

A) IBGP-speaking routers forward BGP updates to all other IBGP-speaking routers within the transit backbone.

B) The IBGP full-mesh requirement within the transit AS is no longer required because BGP routing loops are eliminated.

C) Convergence time within the AS is reduced because fewer routes have to be processed by each IBGP neighbor router.

D) All BGP routers in an AS must be updated directly by the border router receiving the update via EBGP.

Copyright © 2003, Cisco Systems, Inc.

Practice: IBGP Full Mesh

Q1) What are two reasons why a fully meshed logical topology is required for all IBGP neighbors in an AS? (Choose two.)

A) to receive routing information about external networks from border routers

B) to facilitate BGP split-horizon rules

C) to allow IBGP-speaking routers to inject routes into BGP

D) to ensure that each IBGP router has the most current BGP table and attributes

Practice: IBGP Neighbors

Q1) What is the recommended way to run IBGP sessions?

A) Establish IBGP sessions using IP addresses of physical interfaces to allow detection of link failures.

B) Establish an IBGP session using IP addresses of loopback interfaces to prevent IBGP sessions from failing if a physical interface fails.

C) Establish IBGP sessions using the IP addresses of the fastest interface to make sure that packets are forwarded out the fastest interface.

D) There is no recommendation. Any IP address on a router can be used for an IBGP session without any penalties.

Practice: IBGP Next-Hop Processing

Q1) How is the BGP next hop changed inside an AS?

A) The next-hop address is set to the router-ID of the forwarding router.

B) The next-hop address is not modified.

C) The next-hop address is set to the address of the receiver.

D) The next-hop attribute is set to the configured default gateway.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Transit Network Using External Next Hops

Q1) What are the implications of IBGP next-hop processing on the network design?

A) The network must contain a physical full mesh between all BGP-speaking routers.

B) All BGP external routes must be redistributed into the IGP by the border router.

C) All BGP speakers must have IGP reachability to all external neighbors to be able to perform a recursive lookup to resolve next-hop addresses.

D) You must configure static routes on all IBGP-speaking routers pointing to the border routers of the transit AS.

Practice: Transit Network Using Edge Routers as Next Hops

Q1) How can you influence IBGP next-hop processing?

A) by configuring the router to set the next-hop attribute to its router-ID

B) by configuring the router to set the next-hop attribute to the gateway of last resort

C) by configuring the router to ignore the next-hop attribute

D) by configuring the router to set the next-hop attribute to its source address of internal BGP sessions

Practice: Differences Between EBGP and IBGP Sessions

Q1) What are the three major differences that exist between EBGP and IBGP? (Choose three.)

A) No BGP attributes are changed in IBGP updates.

B) Route selection rules slightly prefer IBGP routes.

C) Local preference and MED are propagated only over IBGP sessions.

D) Routes learned from IBGP are not advertised to other IBGP peers.

Packet Forwarding in Transit Autonomous Systems

Practice: Packet Forwarding in a Transit Autonomous System

Q1) Why do you need to run IBGP on all core routers?

A) to ensure that a full mesh exists between all BGP routers in the AS

B) to allow routers to properly resolve the next-hop address

C) to allow routers to forward packets toward all external destinations

D) to ensure correct propagation of the gateway of last resort

Practice: Recursive Lookup in Cisco IOS Software

Q1) What is recursive lookup in Cisco IOS software?

A) verifying that the route exists both in the BGP table and the IP routing table

B) resolving the MAC address of the route to build the MAC header

C) searching the routing table to find a path to the next-hop address

D) the process used to build the IP fast cache on the Cisco IOS router

Practice: Routing Protocols in a Transit Autonomous System

Q1) What are two reasons why you need an IGP in a transit AS? (Choose two.)

A) to provide the gateway of last resort to BGP speakers

B) to perform the recursive lookup for external networks

C) to carry BGP routes through the transit AS

D) to establish IBGP sessions between nondirectly connected routers

 Copyright © 2003, Cisco Systems, Inc.

Practice: BGP and IGP Interaction

Q1) Why should you transport your customer routes in BGP and not in an IGP?

A) to ensure that the gateway of last resort propagates across the AS

B) to correctly resolve addresses in the next-hop attribute

C) to ensure full reachability to external network destinations

D) to protect the IGP from carrying too many routes

Q2) How should BGP react to a failure inside a transit AS that has redundant paths?

A) BGP sessions across the failed link will terminate.

B) BGP does nothing because the IGP finds an alternate path to the neighbors.

C) Failure within the AS causes the BGP to recalculate new paths.

D) BGP will terminate the link between itself and IGP.

Practice: Problems with BGP and IGP Interaction

Q1) What happens when the same route is learned via BGP and an IGP?

A) BGP will install both an EBGP and an IBGP route for the same destination.

B) If the same route is learned via both the IGP and EBGP, the EBGP route is preferred.

C) The route learned through the IGP will be installed in the routing table.

D) If the same route is learned via both the IGP and IBGP, the IBGP route is preferred.

Copyright © 2003, Cisco Systems, Inc.

Configuring a Transit Autonomous System

Practice: Configuring IBGP Neighbors

Q1) Which Cisco IOS command do you use to configure a description on a BGP session?

A) **neighbor** *ip-address* **remote-as** *number* **description** *text*

B) **description** *text*

C) **neighbor** *ip-address* **description** *text*

D) **ip bgp description** *text*

Practice: Configuring IBGP Sessions Between Loopback Interfaces

Q1) Which Cisco IOS command do you use to configure a BGP session between loopback interfaces?

A) **neighbor** *ip-address* **remote-as** *number* **update-source** *interface*

B) **ip bgp source-interface** *interface*

C) **neighbor** *ip-address* **update-source** *interface*

D) **ip bgp update-source** *interface*

Practice: Configuring BGP Synchronization

Q1) What BGP parameter do you need to disable for proper IBGP operation?

A) redistribution

B) AS-path propagation

C) synchronization

D) split horizon

 Copyright © 2003, Cisco Systems, Inc.

Practice: Changing the Administrative Distance of BGP Routes

Q1) How can you change the AD of BGP routes?

A) through AS-path prepending

B) with the **distance bgp** router configuration command

C) by redistributing BGP routes into an IGP

D) by disabling BGP synchronization on the router

Practice: Scalability Limitations of IBGP-Based Transit Backbones

Q1) What are the scalability limitations of an IBGP-based transit AS?

A) The transit backbone can have only two exit points.

B) All EBGP speakers must have a BGP session to all BGP-speaking routers in the transit AS.

C) The transit area is limited to a maximum distance of 100 routers.

D) A full mesh of IBGP sessions is required due to BGP split-horizon rules.

Q2) What two BGP tools can you use to overcome IBGP scalability issues in a transit AS? (Choose two.)

A) disabling synchronization

B) route reflectors

C) BGP confederations

D) BGP peer groups

Monitoring and Troubleshooting IBGP in Transit Autonomous Systems

Practice: Monitoring IBGP

Q1) Which Cisco IOS **show** command indicates that a BGP route is an IBGP route?

A) **show ip route**

B) **show ip route bgp**

C) **show ip bgp**

D) **show ip bgp internal**

Practice: Common IBGP Problems

Q1) What are three common IBGP problems in transit backbones? (Choose three.)

A) IBGP sessions will not become established.

B) IBGP routes are never selected.

C) IBGP routes become “stuck in active state.”

D) BGP routes are not installed into the IP routing table.

Practice: Troubleshooting IBGP Session Startup Issues

Q1) What are three common situations that prevent IBGP sessions from starting? (Choose three.)

A) The IBGP session has been configured to peer to a loopback interface, but **update-source** has not been configured on the neighbor.

B) An access control list filter is blocking access to TCP port 179.

C) The IBGP session has been configured to peer to a loopback interface, but the loopback interface has not been administratively enabled with the **no shutdown** command.

D) The IBGP session has been configured to peer to a loopback interface, but the interfaces are not reachable via the IGP.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Troubleshooting IBGP Route Selection Issues

Q1) What would prevent IBGP routes from being selected as the best route in the BGP table?

A) Failure to disable BGP synchronization.

B) Failure to disable BGP split horizon.

C) The IGP has no route to the BGP next hop.

D) A default route has not been injected into the IGP.

Practice: Troubleshooting IBGP Synchronization Issues

Q1) What common issue could prevent IBGP best routes from being inserted into the IP routing table?

A) Failure to disable BGP synchronization.

B) Failure to disable BGP split horizon.

C) The IGP has no route to the BGP next hop.

D) A default route has not been injected into the IGP.

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

Module 6: Scaling Service Provider Networks

Scaling IGP and BGP in Service Provider Networks

Practice: Common Service Provider Network

Q1) How is routing information exchanged between the service provider and other autonomous systems?

A) IBGP

B) EBGP

C) static routes

D) with an IGP

Q2) In what two ways is routing information typically exchanged between a service provider and its customers? (Choose two.)

A) IBGP

B) EBGP

C) static routes

D) with an IGP

Practice: Route Propagation in Service Provider Networks

Q1) What would be the impact of redistributing BGP into the provider IGP?

A) The transit backbone would stop functioning because all IGPs have smaller administrative distances than IBGP.

B) Convergence times would improve because IGPs have faster convergence and smaller advertisement intervals than BGP.

C) Unbounded, the large number of routes carried in BGP would exceed the scalability of any IGP.

D) There is little to no impact when redistributing BGP routes into the provider IGP.

Copyright © 2003, Cisco Systems, Inc.

Q2) What benefit is gained if links to customer networks are not advertised in the IGP at the provider edge router?

A) IGP performance will increase because the IGP will carry fewer routes.

B) Security is improved because the provider network will hide the customer network.

C) Route summarization is easier because the IGP will carry fewer routes.

D) A flapping link between the customer and the provider will not affect the stability of the provider IGP.

Practice: Scaling Service Provider Routing Protocols

Q1) What are the two responsibilities of the IGP? (Choose two.)

A) carrying internal networks for optimal routing

B) BGP next-hop resolution

C) route propagation to external autonomous systems

D) exchanging routing information with customer networks

Q2) What are the responsibilities of BGP? (Choose two.)

A) advertising router loopback interfaces for IBGP sessions

B) propagation of customer routes

C) resolving BGP next-hop addresses

D) exchanging routing information with other providers

 Copyright © 2003, Cisco Systems, Inc.

Practice: Scaling Service Provider Addressing

Q1) What is the impact of using private addresses on the internal links of the service provider backbone?

A) If the provider uses private addressing, the customer must use public addressing.

B) A traceroute run from inside a customer network protected with a firewall is broken.

C) Private addressing prohibits optimal path selection in BGP.

D) If private addressing is used in the service provider backbone, route summarization must be disabled.

Q2) Why should loopback interface addresses not be summarized?

A) Summarizing the loopback interfaces will cause IBGP sessions to become inactive.

B) If the loopback interfaces are summarized, they will be unreachable from other routers in the POP.

C) Summarized loopback interfaces might interfere with optimal routing between loopback interfaces used for IBGP peering sessions.

D) Loopback interfaces can be summarized without any adverse effects.

Copyright © 2003, Cisco Systems, Inc.

Introduction to Route Reflectors

Practice: IBGP Scalability Issues in a Transit AS

Q1) What is a scalability limitation of IBGP-based transit autonomous systems?

A) A single AS is limited to 100 routers.

B) A full mesh of IBGP sessions is required due to BGP split-horizon rules.

C) Only one instance of BGP is allowed in a single BGP domain.

D) Each BGP AS is limited to a single IGP.

Practice: Route Reflector Split-Horizon Rules

Q1) Which three types of routes does a route reflector propagate to its clients? (Choose three.)

A) EBGP

B) client IBGP

C) nonclient IBGP

D) IGP

Q2) Which routes are propagated by a route reflector client to its IBGP neighbors?

A) EBGP

B) client IBGP

C) nonclient IBGP

D) IGP

 Copyright © 2003, Cisco Systems, Inc.

Practice: Redundant Route Reflectors

Q1) What are a benefit and a drawback of redundant route reflector designs? (Choose two.)

A) Each client will receive only a single copy of BGP route update information.

B) IBGP routing loops can occur.

C) The transit AS does not need an IGP.

D) Redundancy eliminates a single point of failure in the route reflector.

Practice: Route Reflector Clusters

Q1) What are route reflector clusters used for?

A) to detect and correct route instability

B) to differentiate between route reflector clients and nonclients

C) to eliminate EBGP split-horizon rules

D) to detect potential routing loops in redundant configurations

Practice: Additional Route Reflector Loop Prevention Mechanisms

Q1) What two loop-prevention mechanisms support route reflectors? (Choose two.)

A) cluster-list

B) originator-ID

C) AS-path attribute

D) spanning tree

Copyright © 2003, Cisco Systems, Inc.

Network Design with Route Reflectors

Practice: Network Design with Route Reflectors

Q1) With what should a route reflector client have IBGP sessions?

A) with other clients only

B) with its closest neighbor EBGP peer only

C) with its route reflectors and directly connected EBGP peers

D) with its route reflectors and other clients in the cluster

Practice: Potential Network Issues

Q1) What happens if a route reflector client establishes sessions with route reflectors in two clusters?

A) This is an invalid configuration.

B) It introduces the originator-ID attribute and the potential for routing loops.

C) BGP updates from the client will not be forwarded into the full mesh with different cluster-IDs.

D) Some routes will be reflected by more than one cluster, resulting in increased size of the BGP table for route reflectors and for the offending client.

Q2) What happens if two route reflectors have each other configured as a client?

A) This is an invalid configuration.

B) Routing loops will occur because the cluster-list will not be updated properly.

C) The only result will be increased BGP traffic.

D) BGP updates from the client will not be forwarded into the full mesh with different cluster-IDs.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Hierarchical Route Reflectors

Q1) What is a hierarchical route reflector design?

A) a design where there is only one route reflector per cluster

B) a design where route reflectors are located in distinct, nonoverlapping clusters

C) a design where some route reflectors are also clients of other route reflectors

D) a design with multiple route reflectors per cluster

Copyright © 2003, Cisco Systems, Inc.

Configuring and Monitoring Route Reflectors

Practice: Route Reflector Backbone Migration

Q1) Which two BGP parameters do you have to configure on a route reflector? (Choose two.)

A) cluster-ID

B) originator-ID

C) cluster-list

D) route reflector clients

Q2) Which BGP parameter do you have to configure on a route reflector client?

A) cluster-ID

B) originator-ID

C) route reflector identification

D) No configuration is required on the client routers.

Practice: Configuring Route Reflectors

Q1) What are three migration steps required to convert from a fully meshed IBGP AS to an AS based on route reflectors? (Choose three.)

A) Remove unnecessary IBGP sessions.

B) Configure the clients on the route reflectors.

C) Configure IBGP sessions between route reflector clients.

D) Configure the cluster-ID on the route reflectors.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Monitoring Route Reflectors

Q1) Which command should you use to display the cluster-list attribute, identifying whether an IBGP route was reflected?

A) **show ip bgp** *prefix*

B) **show ip bgp** *cluster*

C) **show ip bgp** *interface*

D) **show ip bgp** *attribute-id*

Q2) Which command should you use to identify route reflector clients without inspecting the router configuration?

A) **show ip bgp** *prefix*

B) **show ip bgp neighbors**

C) **show ip bgp clients**

D) **show ip bgp summary**

Introduction to Confederations

Practice: IBGP Transit AS Problems

Q1) How do BGP confederations reduce the IBGP full mesh?

A) Confederations allow direct peering between external routers through a transit AS.

B) Confederations break an AS up into small “mini” autonomous systems forming EBGP sessions within the confederation.

C) Confederations allow a router to advertise IBGP-learned routing information to other IBGP speakers, reducing the total number of peers required.

D) Confederations combine external autonomous systems into a larger “master” AS treating all internal routers as EBGP peers.

Practice: Splitting a Transit AS with BGP Confederations

Q1) What IBGP rule applies to member autonomous systems within a confederation?

A) The AS number in each member-AS must be from within the private AS number range.

B) Each member-AS within the confederation must use a globally registered AS number.

C) Each member-AS must be a transit AS for the confederation.

D) Routers contained within the member-AS must be fully meshed.

 Copyright © 2003, Cisco Systems, Inc.

Practice: AS-Path Propagation Within the BGP Confederation

Q1) How is the AS path propagated between intraconfederation EBGP peers?

A) The confederation AS number is prepended to the AS-path attribute to ensure that the AS path is valid when the route update is sent to the EBGP peer in another AS.

B) Intraconfederation EBGP peers maintain a second AS path so that IBGP mesh requirements can be maintained without affecting the AS path used to reach EBGP peers.

C) The member-AS number is prepended to the AS-path attribute.

D) Because the confederation resides in a single AS, IBGP rules apply, and no AS number is prepended to the AS-path attribute.

Q2) What happens to the AS path when a BGP update is sent to an external peer outside of the confederation?

A) The AS path is not modified because it already contains the AS numbers of each member-AS.

B) The AS number of the confederation is prepended to the existing AS path.

C) The intraconfederation AS numbers are removed from the path, and the external AS is prepended.

D) All member-AS entries are removed from the AS-path attribute by the EBGP peer.

Practice: AS-Path Processing in BGP Confederations

Q1) What happens if a router not supporting BGP confederations is placed inside the confederation?

A) The router will learn about the confederation through BGP capabilities exchange.

B) The router will process BGP updates as normal because it has to be aware only of the member-AS to which it belongs.

C) The router will automatically convert the intraconfederation AS numbers to the external AS number of the confederation.

D) The router will be unable to interpret the AS-path attribute and terminate its BGP session with that peer.

Copyright © 2003, Cisco Systems, Inc.

Practice: Intraconfederation EBGP Session Properties

Q1) What is the difference between an EBGP session and intraconfederation EBGP session?

A) EBGP and intraconfederation EBGP sessions are identical in how they treat EBGP peers.

B) EBGP sessions do not allow sessions to peers that are not directly attached.

C) Intraconfederation EBGP routers retain IBGP processing rules of the local preference, MED, and next-hop attributes.

D) Intraconfederation EBGP routers can form BGP peering relationships only between member-AS routers.

 Copyright © 2003, Cisco Systems, Inc.

Configuring and Monitoring Confederations

Practice: BGP Confederation Design Rules

Q1) How can you reduce the IBGP full mesh within a confederation AS?

A) You cannot reduce the full mesh because all IBGP peers must be fully meshed within the member-AS.

B) Implementing router reflectors inside a member-AS can reduce the IBGP full-mesh requirement.

C) You can nest a confederation within a confederation to remove the IBGP full-mesh requirement.

D) Because confederations are used, there is no requirement for an IBGP full mesh within each member-AS.

Practice: Planning BGP Confederations

Q1) Which three requirements are BGP confederation planning steps? (Choose three.)

A) Define an AS number for each member-AS.

B) Remove the IBGP mesh in each member-AS.

C) Divide the transit AS into smaller autonomous systems.

D) Verify that the Cisco IOS version installed on the router supports confederations.

Practice: Configuring BGP Confederations

Q1) Which two BGP parameters do you need to specify on every router within a confederation? (Choose two.)

A) a list of all AS numbers in the confederation

B) a list of all true EBGP sessions

C) the official AS number (as the identifier of the confederation)

D) the correct MD5 authentication password in each peer

Copyright © 2003, Cisco Systems, Inc.

Q2) What impact will configuring BGP confederations have on the BGP network?

A) BGP update performance will be increased because the member-AS mesh is reduced.

B) EBGP peers will have a reduced number of **neighbor** statements to connect to the transit backbone.

C) Network availability will be impacted because the BGP routing process must be removed and restored using the new AS numbers.

D) IBGP session establishment times will be reduced.

Q3) What does the BGP confederation identifier define?

A) the AS number of the confederation external peer

B) the public AS number that the confederation is using externally

C) the AS number of the member-AS

D) the MD5 authentication password for the confederation

Practice: Monitoring BGP Confederations

Q1) How will the **show ip bgp** command display the intraconfederation segment of the AS path?

A) As a regular entry in the AS-path attribute.

B) As a separate AS-path list independent of the AS-path attribute.

C) As an entry in the AS path designated by parentheses.

D) The intraconfederation AS is not displayed because it is not a part of the EBGP AS path.

Q2) How does the **show ip bgp** command indicate that an EBGP session is with an intraconfederation peer instead of an actual EBGP peer?

A) by indicating that the neighbor is under common administration

B) by specifying that the session is over a confederation internal link

C) by displaying the next-hop address in parentheses

D) by indicating the BGP established state of the session to the external neighbor

 Copyright © 2003, Cisco Systems, Inc.

Module 7: Optimizing BGP Scalability

Improving BGP Convergence

Practice: BGP Convergence

Q1) Why is the speed at which a routing protocol converges important?

A) A nonconverged routing protocol cannot forward packets.

B) Routing metrics cannot be properly computed unless the network is converged.

C) Routing algorithms that converge slowly can cause routing loops or network outages.

D) Networks that are not converged have reduced security.

Practice: BGP Processes

Q1) What two BGP processes can consume large amounts of CPU resources? (Choose two.)

A) BGP open

B) BGP scanner

C) BGP I/O

D) BGP router

Practice: CPU Effects of BGP Processes

Q1) Why are BGP filtering tools such as AS-path filters and prefix-lists important in managing CPU resources in BGP-enabled routers?

A) Because filtering tools make the router inherently more secure.

B) These tools help reduce the size of the BGP table that is walked by BGP processes.

C) These tools switch BGP processing into CEF switching mode.

D) These tools reduce the number of BGP attributes attached to BGP routes.

Copyright © 2003, Cisco Systems, Inc.

Practice: Improving BGP Convergence

Q1) What BGP convergence-enhancing tool reduces convergence by reducing transport overhead and processing delays at network transit points?

A) aggressive TCP packet queuing

B) PMTU discovery

C) scan time

D) peer groups

Q2) What two BGP convergence-enhancing tools are recommended for all BGP implementations? (Choose two.)

A) PMTU discovery

B) scan time

C) advertisement interval

D) peer groups

Practice: Path MTU Discovery

Q1) In what two ways does the PMTU discovery feature aid in improving BGP convergence? (Choose two.)

A) It allows the router to determine the MSS of the network path.

B) It allows the router to only send TCP packets only across network paths with the highest MSS.

C) It reduces the amount of overhead in transporting data between TCP sessions.

D) It removes the increased CPU and memory overhead required to fragment an IP datagram.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Increasing Input Queue Depth

Q1) How does increasing the size of the input queue on an interface aid in improving BGP convergence?

A) It allows for larger packets to be sent between BGP neighbors.

B) It increases the number of BGP updates that can be contained in a single TCP segment.

C) It reduces overhead in processing BGP update packets by reducing the number of retransmissions caused by dropped TCP ACK packets.

D) It increases the rate at which incoming packets can be sent to the output queue of the outgoing interface.

Practice: BGP Scan Time

Q1) What negative effects can occur when you are reducing the scan time in an attempt to improve BGP convergence?

A) Frequent runs of the BGP scanner process may cause BGP to converge before the IGP does, resulting in network black holes.

B) If the scanner process runs too frequently, the router CPU and memory resources may become overwhelmed.

C) If the BGP scanner process runs too often, the BGP router process will not get enough CPU cycles to properly update the routing table.

D) If the BGP scanner process runs too often, the BGP open process will not get enough CPU cycles to properly maintain BGP neighbor relationships.

Copyright © 2003, Cisco Systems, Inc.

Practice: BGP Advertisement Interval

Q1) How does the advertisement interval timer act as a rate-limiting feature in BGP?

A) The advertisement interval timer specifies the time between successive runs of the BGP scanner process, resulting in rate-limiting the update process.

B) Adjusting the advertisement interval limits the rate at which the router sends and receives route advertisements.

C) Adjusting the advertisement interval does not limit the rate of route selection but only the rate of route advertisement.

D) Adjusting the advertisement interval limits the rate at which the router can receive BGP updates from its neighbors.

 Copyright © 2003, Cisco Systems, Inc.

Limiting the Number of Prefixes Received from a BGP Neighbor

Practice: Limiting the Number of Routes Received from a Neighbor

Q1) What are three reasons to limit the number of BGP prefixes received from a neighbor? (Choose three.)

A) to prevent denial-of-service attacks

B) to protect against incorrect router configuration on the neighbor side

C) to prevent redundant routing information from being loaded into the BGP table

D) to avoid overloading router memory and CPU resources

Practice: Configuring the BGP Maximum-Prefix Function

Q1) By default, what happens when the number of configured prefixes is exceeded?

A) A filter is automatically generated and applied to incoming BGP updates.

B) A syslog message is sent to the log.

C) The BGP session is restarted to clear out the BGP table.

D) The BGP session is terminated and will remain inactive until the session is manually enabled once again.

Practice: Monitoring the BGP Maximum-Prefix Function

Q1) In what two situations would a directly connected BGP neighbor stay in the Idle state? (Choose two.)

A) The neighbor has exceeded the maximum number of allowed prefixes.

B) The maximum-prefix threshold has been reached.

C) The **restart** option has not been specified with the **maximum-prefix** command.

D) The neighbor is more than one hop away.

Copyright © 2003, Cisco Systems, Inc.

BGP Peer Groups

Practice: Peer Group Requirements

Q1) What is the purpose of using BGP peer groups?

A) BGP peer groups can be used to configure the same set of parameters for a number of BGP neighbors in a common template.

B) BGP peer groups can be used to allow anonymous BGP neighbors.

C) BGP peer groups allow EBGP peers to be configured with the same AS number and parameters.

D) BGP peer groups can be used to hide the identity of BGP peers from external neighbors.

Practice: Peer Groups as a BGP Performance Tool

Q1) What are two performance benefits of configuring BPG peer groups on a Cisco router? (Choose two.)

A) Peer groups reduce the amount of memory required to process BGP updates between BGP peers.

B) Peer groups reduce the amount of CPU resources required to process many BGP neighbor sessions.

C) Peer groups remove route replication requirements when routers process updates from a large number of neighbors.

D) Peer groups reduce the number of TCP sessions required to form BGP neighbor relationships.

Practice: BGP Peer Group Limitations

Q1) What are two limitations of BGP peer groups on Cisco routers? (Choose two.)

A) EBGP and IBGP neighbors cannot be members of the same peer group.

B) All routers in the peer group must belong to the same AS.

C) Peer group members cannot contain different outbound filtering mechanisms.

D) Peer group members must have the same inbound filtering mechanisms.

 Copyright © 2003, Cisco Systems, Inc.

Practice: BGP Peer Groups in Cisco IOS Software

Q1) Why must outgoing filters be applied uniformly to all members of a BGP peer group?

A) Outgoing filters are applied only once to each peer group member to conserve router CPU resources.

B) Outgoing filters are applied to the leader of the peer group first, and the results are replicated to other peer group members.

C) Outgoing filters are limited to a specific router-processing path that prevents them from being applied differently to different neighbors.

D) There is no restriction on applying outgoing filters to members of a peer group.

Practice: Configuring Peer Groups

Q1) What are three steps required to properly configure BGP peer groups on Cisco routers? (Choose three.)

A) Specify parameters for the BGP peer group.

B) Create a BGP peer group.

C) Enable the peer group by clearing the BGP session.

D) Assign a neighbor into the peer group.

Q2) What happens if a neighboring router that is configured as a member of a peer group does not support BGP peer groups?

A) The BGP peer group negotiation fails, and the neighbor is removed from the peer group.

B) The BGP session negotiation fails; it oscillates between Idle and Active states.

C) The local router will reject the peer group configuration.

D) The BGP peer group functionality does not require any support by the neighbor.

Copyright © 2003, Cisco Systems, Inc.

Practice: Monitoring Peer Groups

Q1) What command do you use to display the summary status of all neighbors in a peer group?

A) **show ip bgp**

B) **show peer-group summary**

C) **show ip bgp neighbor**

D) **show ip bgp peer-group summary**

 Copyright © 2003, Cisco Systems, Inc.

BGP Route Dampening

Practice: BGP Route Dampening

Q1) What two mechanisms are built into BGP to make it more scalable by reducing the route-processing requirements of BGP routers? (Choose two.)

A) split horizon

B) route dampening

C) synchronization

D) per-neighbor update timers

Practice: Route-Dampening Operation

Q1) What are two things that happen to an EBGP route that has become unreachable when BGP route dampening is used? (Choose two.)

A) It is removed from the IP routing table.

B) It is removed from the BGP table.

C) It will remain in the IP routing table as long as its penalty remains greater than 50 percent of the reuse limit.

D) It is kept in the BGP table and marked as a history entry.

Q2) What type of routes can be dampened?

A) IGP

B) IBGP

C) EBGP

D) locally advertised routes

Copyright © 2003, Cisco Systems, Inc.

Practice: Configuring BGP Route Dampening

Q1) What are two methods of enabling route dampening on a Cisco router? (Choose two.)

A) globally, by enabling route dampening in global router configuration mode

B) globally, by enabling route dampening under the BGP routing process

C) on specific routes by enabling route dampening on a specific interface

D) by using a route-map in the BGP process to apply route dampening to specific routes.

Q2) Given a reuse of 1000 and a suppress of 3000, how should you configure the half-life parameter if a route flapping four times is reusable after 20 minutes?

A) 5 min

B) 10 min

C) 15 min

D) 20 min

Practice: Releasing Dampened Routes

Q1) A network has been suppressed due to route dampening. What are two things that you can do on the router to unsuppress the dampened route? (Choose two.)

A) Clear the flap statistics for the dampened prefix.

B) Issue a **clear ip bgp dampening** command on the router.

C) Wait for the route penalty to fall below the reuse limit.

D) Reset the BGP process on the neighbor advertising the route.

 Copyright © 2003, Cisco Systems, Inc.

Practice: Monitoring Route Dampening

Q1) What two things could happen to a BGP route that is penalized above the reuse limit, but its assigned penalty is under the suppress limit? (Choose two.)

A) The route is suppressed from BGP updates if it is reachable.

B) The route is marked as a history entry in the BGP table.

C) The route is withdrawn from the IP routing table.

D) The route will continue to be advertised.

Q2) When you are using default route-dampening parameter values, what happens to a suppressed route after its penalty decays for a time period equal to the half-life?

A) The route is released and available as a BGP candidate for best-path selection.

B) The route flap statistics are cleared, but the route will remain suppressed.

C) The route will remain suppressed for at least one more half-life.

D) The route will be released provided that no additional flaps have occurred.

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

Lesson Assessment Solutions

The lesson assessment items and solutions are listed here. The shaded answer options are the correct answers.

Module 1: BGP Overview

Introduction to BGP

Quiz: Introduction to BGP

Q1) What three items are BGP enhancements to traditional distance-vector routing protocols? (Choose three.)

A) reliable updates

B) use of triggered updates only

C) enhanced security

D) rich metrics

Q2) What protocol facilitates reliable update capabilities in BGP?

A) TCP

B) UDP

C) HSRP

D) ICMP

Copyright © 2003, Cisco Systems, Inc.

Q3) What are three characteristics of an AS? (Choose three.)

A) uses IGPs for intradomain routing

B) uses EGPs for interdomain routing

C) is a collection of networks under a common administrative authority

D) consists of a group of network domains

Q4) What three scenarios are common scenarios where BGP is used? (Choose three.)

A) a customer with a connection to multiple service providers

B) service provider networks acting as transit systems and forwarding external traffic through their network

C) single-site customer intranet with complex administrative policies between departments

D) as the core routing protocol in very large enterprise networks

Q5) What are three recommended BGP use guidelines for multihomed customer networks? (Choose three.)

A) Most multihomed customers should use BGP with their service providers.

B) Most multihomed customers should forward routing information received from one provider to the other provider.

C) The multihomed customer must have its own public AS number.

D) Multihomed customers should use a provider-independent, public address space.

Q6) What is a limitation of the BGP routing protocol?

A) You cannot use BGP to implement hop-by-hop routing policy controls.

B) You cannot use BGP to influence the routing policy in a downstream AS.

C) BGP cannot control forwarding of packets based on their destination address.

D) BGP cannot scale to very large networks with more than 110,000 routes.

 Copyright © 2003, Cisco Systems, Inc.

BGP Session Establishment

Quiz: BGP Session Establishment

Q1) What is indicated by a state of "Idle" in the output of the **show ip bgp summary** command?

A) The router is currently not attempting to establish a connection with a neighbor.

B) The connection to the configured neighbor has timed out.

C) The connection to a BGP neighbor has been established, and no errors have been received on the connection.

D) The connection to a BGP neighbor has been established, and no packets have been sent.

Q2) What happens if two TCP connection attempts between configured BGP neighbors succeed?

A) Both connections will be terminated, and the neighbors will re-establish a neighbor relationship.

B) One connection will be maintained as primary and the other as backup.

C) One of the two connections will be torn down.

D) The router with the lower router-ID will determine if the second connection is torn down or used as a backup TCP connection.

Q3) Given the following BGP session states:

1. OpenConfirm
2. Established
3. Idle
4. OpenSent
5. Active

What is their order of progression during the creation of a successful neighbor session?

A) 5, 1, 4, 2, 3

B) 3, 4, 1, 5, 2

C) 5, 4, 1, 3, 2

D) 3, 5, 4, 1, 2

Copyright © 2003, Cisco Systems, Inc.

Q4) What does the field “TblVer” indicate in the output of the **show ip bgp summary** command?

A) the current version of BGP in use by the router

B) the number of route prefixes contained in the BGP update of the router

C) BGP messages received from that neighbor

D) last version of the BGP database that was sent to that neighbor

Q5) What occurs when you use MD5 between two BGP neighbors?

A) Every packet is encrypted with MD5.

B) The IP header is encrypted using MD5.

C) An MD5 checksum is calculated and sent with each packet so that its source can be verified.

D) A username and password is embedded in an IP datagram that is matched to a username and password on the remote neighbor.

 Copyright © 2003, Cisco Systems, Inc.

BGP Route Processing

Quiz: BGP Route Processing

Q1) What does a router running BGP do with a BGP update containing its own AS path?

A) The router checks to see if the information contained in the update is better than its current information. If it is, it will update its BGP table.

B) The router accepts the route update.

C) The router silently discards (denies) the route.

D) The router will return an error to the router sending the update.

Q2) How many alternate paths to a single destination will a BGP router maintain in the BGP table?

A) The router will maintain only the best path to the destination.

B) The router will maintain two paths, the best path and a backup route.

C) The BGP table will hold up to four routes by default and a maximum of six configurable routes.

D) The BGP table will store all valid, advertised routes to the destination in the BGP table.

Q3) What are two ways in which local networks are advertised into the BGP routing protocol process? (Choose two.)

A) automatically, after a BGP neighbor session is established

B) manually, with the **network** command

C) through redistribution into the BGP process

D) by advertising them to the BGP table on the router after CDP discovers connected networks

Copyright © 2003, Cisco Systems, Inc.

Q4) What are two situations when is it appropriate to disable automatic summarization in BGP? (Choose two.)

A) when BGP neighbors are not configured to advertise aggregate routes to upstream providers

B) when the classless variant of the **network** command is used

C) when you are using a classless IGP in the AS

D) when the effects of automatic summarization of IGP-to-BGP redistribution are not desired

Q5) What is the AD of BGP routes in the IP routing table learned from BGP neighbors in a different AS?

A) 1

B) 20

C) 90

D) 120

Q6) What three BGP attributes are displayed for each route in the BGP table when you are using the **show ip bgp** command? (Choose three.)

A) weight

B) communities

C) origin

D) AS-path

 Copyright © 2003, Cisco Systems, Inc.

Basic BGP Configuration

Quiz: Basic BGP Configuration

Q1) What is the valid AS number range for a BGP process on a Cisco router?

A) 1 – 256

B) 1 – 32768

C) 1 – 65535

D) 1 – 131072

Q2) What AS numbers are defined as private AS numbers?

A) 1 – 128

B) 32768 – 64511

C) 64512 – 65535

D) 65536 – 131072

Q3) What two parameters must you configure with the **neighbor** command to establish a BGP session with an external neighbor? (Choose two.)

A) neighbor IP address

B) subnet mask of the IP network

C) remote AS number

D) local AS number

E) description of the neighbor

Q4) What is the best method to temporarily disable a BGP neighbor session?

A) Remove the **neighbor** command from the BGP router process.

B) Remove the BGP router process from the configuration.

C) Terminate the neighbor connection with the **neighbor shutdown** command.

D) Disconnect the neighbor by initiating a router reload.

Copyright © 2003, Cisco Systems, Inc.

Q5) What three steps must you complete to advertise a classless prefix into BGP? (Choose three.)

A) Configure the prefix with the **network** command.

B) Specify the **mask** keyword with the locally advertised route.

C) Configure **redistribute connected** under the BGP router process.

D) Use a static route pointing to null 0 that matches the prefix.

Q6) What origin code is carried with routes redistributed into BGP?

A) internal

B) external

C) unknown

D) incomplete

Q7) What must be true for a BGP route aggregate to be advertised in the IP routing table?

A) At least one network in the specified range must exist in the BGP table.

B) You must configure automatic summarization under the BGP routing protocol process.

C) You must configure a route to null 0 matching the aggregate.

D) No synchronization must be configured under the BGP routing protocol process.

Q8) What are two benefits of using route aggregation in BGP? (Choose two.)

A) It ensures that even if aggregate networks are down, the aggregate is advertised, eliminating “black holes.”

B) It reduces the amount of memory used in the router to store the BGP table.

C) It reduces route flapping and its effects on router CPU resources.

D) BGP attribute granularity is maintained, ensuring optimal path selection.

 Copyright © 2003, Cisco Systems, Inc.

Module 2: Route Selection Using Policy Controls

Outbound Route Filtering

Quiz: Outbound Route Filtering

Q1) What best describes the capabilities of the proprietary ORF type supported on Cisco routers?

A) standard BGP communities filtering

B) extended BGP communities filtering

C) AS-path filtering

D) prefix-list filtering

Q2) What are two key benefits to using outbound route filtering? (Choose two.)

A) conserves CPU cycles

B) improves security

C) reduces bandwidth used by unnecessary routing updates

D) increases neighbor availability

Q3) How should you configure the **neighbor capability orf prefix-list** on a router applying a prefix-list filter as an outbound route policy?

A) send

B) receive

C) both

D) prefix-filter

Copyright © 2003, Cisco Systems, Inc.

Q4) What are two methods of determining that a router has ORF capabilities exchange configured? (Choose two.)

A) with the **show running-config | begin bgp** command

B) using the **show ip bgp negotiate** command

C) by executing the **show ip bgp neighbors** command

D) by using the **show ip prefix-list** command

Q5) What are two prerequisites before you can configure ORF prefix-list functionality? (Choose two.)

A) A route refresh must be sent using the **clear ip bgp** command.

B) A BGP peering session between the ORF routers must be up and running.

C) ORF capabilities must be enabled on both routers.

D) You must configure a prefix-list filter on the receiving router.

 Copyright © 2003, Cisco Systems, Inc.

Route-Maps as BGP Filters

Quiz: Route-Maps as BGP Filters

Q1) What are three commonly used route-map match criteria in BGP environments? (Choose three)

A) AS-path

B) prefix-list

C) community attribute

D) local preference

Q2) How do you implement a “permit all” when you are using route-maps?

A) By default, a route-map has an “implicit permit any” if no match is found.

B) You must configure a route-map with a “permit” parameter and no match clause.

C) You must configure a route-map with a “deny” parameter and a “deny none” clause.

D) You must configure a route-map with a “permit any” match clause.

Q3) What happens to incoming BGP updates that do not match any route-map match clauses?

A) They are entered into the BGP table.

B) They are entered into the BGP table and marked with a weight of 32768.

C) They are not accepted by the router or entered into the BGP table.

D) They are entered into the BGP table if a matching route exists in the IP routing table.

Q4) What three BGP attributes can you set using route-maps? (Choose three.)

A) MED

B) weight metric

C) next-hop

D) atomic aggregate

Copyright © 2003, Cisco Systems, Inc.

Q5) What are three uses of route-maps in a BGP environment? (Choose three.)

A) to filter incoming prefixes based on a prefix and the AS-path attribute

B) to modify routing information currently in the BGP table

C) to set BGP attributes such as weight and local preference on outgoing updates

D) to filter the redistribution of IGP routes into BGP

Q6) What are two reasons for using route-map sequence numbers? (Choose two.)

A) to allow insertion or deletion of route-map entries

B) to order the execution sequence of route-map match clauses

C) to provide an ordered execution sequence for the route-map

D) to map between prefix-list statements and route-map match clauses

 Copyright © 2003, Cisco Systems, Inc.

Module 3: Route Selection Using Attributes

BGP Path Attributes

Quiz: BGP Path Attributes

Q1) Which three statements are true of BGP mandatory well-known attributes? (Choose three.)

A) They must be present in all BGP updates.

B) All BGP-compliant implementations must recognize them.

C) All BGP-compliant routers must adhere to policies specified in mandatory attributes.

D) All well-known attributes are propagated to other neighbors.

Q2) Which three attributes are BGP mandatory well-known attributes? (Choose three.)

A) next-hop

B) weight

C) AS-path

D) origin

Q3) What three possible values are assigned to the BGP origin attribute? (Choose three.)

A) IGP

B) EGP

C) unknown

D) internal

Q4) What nontransitive optional BGP attribute is useful in assisting with the route selection process when multiple links to another AS exist?

A) next-hop

B) local preference

C) MED

D) AS-path

Copyright © 2003, Cisco Systems, Inc.

Q5) How is the BGP next-hop attribute modified?

A) If the next-hop attribute is in the same IP subnet as the receiving router, the attribute is unchanged; otherwise, it is set to the IP address of the sending router.

B) The next-hop attribute is always set to the IP address of the sending router.

C) The next-hop attribute is modified only when BGP packets exit an AS.

D) The BGP next-hop attribute is modified only when BGP packets traverse point-to-point links.

Q6) Which three statements are true regarding the BGP AS-path attribute? (Choose three.)

A) The local AS number is prepended to the AS path each time that the route crosses an AS boundary.

B) The AS that originally injected the route into BGP is always found in the rightmost end of the AS path.

C) The AS-path attribute is also used to avoid routing loops.

D) BGP routes with an empty AS path were injected into BGP from outside the local AS.

 Copyright © 2003, Cisco Systems, Inc.

AS-Path Prepending

Quiz: AS-Path Prepending

Q1) What is AS-path prepending?

A) when a router, sending a BGP update, adds the AS number of the router from which it received the route, to the AS-path attribute.

B) when a router, sending a BGP update, adds the AS number of the router to which it is sending the route, to the AS-path attribute.

C) when a router, sending a BGP update, adds its AS number to the AS-path attribute multiple times

D) when a router uses the AS-path attribute in route selection

Q2) AS path will be the route selection criterion used when which of the following is true?

A) It is the first criterion used in BGP route selection.

B) When there is no difference in weight, local preference, or route origination.

C) When the multi-exit discriminator is identical on the candidate routes.

D) The weight, local preference, MED, and origin attributes must be identical before the AS-path attribute is used for route selection.

Q3) What command do you use to manipulate the AS-path attribute?

A) the global configuration command, **set as-path prepend** *as-number*

B) the router configuration command, **set as-path prepend** *as-number*

C) **set as-path prepend** *as-number* in a route-map

D) the interface global command, **set as-path prepend** *as-number*

Copyright © 2003, Cisco Systems, Inc.

Q4) Given the following configuration from a router in AS 347, advertising network 11.0.0.0/8 to an EBGP neighbor 2.0.0.2 in AS 529:

```
route-map addAS permit 10
       set as-path prepend 347 347 347

router bgp 347
       neighbor 2.0.0.2 remote-as 529
       neighbor 2.0.0.2 route-map addAS out
```

What are the contents of the AS-path attribute for route 11.0.0.0/8 on a router residing in AS 529?

A) 347 347 347

B) 347 347 347 347

C) 529 347 347 347

D) 529 347 347 347 347

Q5) Why do network administrators need to use AS-path prepending?

A) AS-path prepending allows a customer to potentially influence return path route selection.

B) AS-path prepending is used on a customer router to control outgoing route updates.

C) Service providers use AS-path prepending to control incoming updates from a customer AS.

D) AS-path prepending is used between service providers who are both connected to a customer AS, to determine who will be the primary link to the customer.

Q6) How does AS-path prepending affect a router?

A) AS-path prepending is simply a term used to describe when a router uses the AS-path attribute in route selection and hence does not affect router resources.

B) The longer the AS-path attribute attached to BGP updates, the more router memory requirements increase.

C) AS-path prepending does not impact the router because Cisco IOS software recognizes that AS-path prepending is in use and stores a single AS number with a pointer to the number of AS-path prepends.

D) AS-path prepending causes the router to operate in process-switching mode because the BGP update must be stored, manipulated, and then rewritten to accommodate for the new AS-path attribute.

 Copyright © 2003, Cisco Systems, Inc.

Module 4: Customer-to-Provider Connectivity with BGP

Customer-to-Provider Connectivity Requirements

Quiz: Customer to Provider Connectivity Requirements

Q1) If a customer required additional bandwidth as well as redundancy, what method would be preferred?

A) a single permanent connection to one ISP

B) permanent connections to more than one ISP

C) dial-up connections to more than one ISP

D) multiple permanent connections to one ISP

Q2) What type of redundancy do multiple permanent connections providing load-sharing configuration display?

A) link

B) equipment

C) service provider

D) routing protocol

Q3) In a customer-to-provider routing scheme, what method of routing is preferred due to its lower complexity?

A) policy-based routing

B) dynamic routing

C) content routing

D) static routing

Copyright © 2003, Cisco Systems, Inc.

Q4) Why is it that with multiple permanent connections to more than one ISP, the use of dynamic routing with BGP is required?

A) When one of the connections is lost, the link level detects this loss and places the interface in a down state.

B) Monitoring the link status cannot detect a problem inside one of the ISP networks.

C) Static routes detect problems inside one of the ISP networks.

D) It is not required, and static routing may be used.

Q5) What can be done when a customer is assigned only a very small subnet of public addresses?

A) Purchase more addresses as required.

B) Use NAT.

C) Add a service provider.

D) Add links to the same service provider.

Q6) What are two different addressing schemes that customers use to connect to a service provider? (Choose two.)

A) provider-independent

B) customer-independent

C) provider-assigned

D) customer-assigned

Q7) Which two of the following criteria are required for a customer to be multihomed to multiple ISPs? (Choose two.)

A) A public AS number.

B) A private AS number.

C) The customer must run BGP with both of its ISPs.

D) The customer must run BGP with one ISP and may use static routing with the other.

 Copyright © 2003, Cisco Systems, Inc.

Implementing Customer Connectivity Using Static Routing

Quiz: Implementing Customer Connectivity Using Static Routing

Q1) What are two circumstances where you can use static routing as part of installing redundant connections between the customer network and a single service provider network? (Choose two.)

A) The router must be able to detect a link failure.

B) The default route must be announced using the customer IGP.

C) If one link goes down, the interface must remain in an up state.

D) The customer IGP must continue to advertise the static default route.

Q2) A customer route that should not be announced to the rest of the Internet is marked using what?

A) a route tag

B) the **export** community

C) the **no-export** community

D) he public address filter

Q3) When you are designing static route propagation in a service provider network, what three steps must you take? (Choose three.)

A) Assign a tag to each combination of services.

B) Configure a community that matches defined tags.

C) Redistribute static routes into BGP through a route-map.

D) Identify all possible combinations of services offered to a customer.

Q4) What does a route-map assign that will be used by other routers within a network?

A) a tag

B) community values

C) public addressing

D) QoS

Copyright © 2003, Cisco Systems, Inc.

Q5) What three key pieces of information can you derive from the following router command output? (Choose three.)

```
AS387_Backup# sh ip bgp 11.2.3.0
BGP routing table entry for 11.2.3.0/24, version 7
Paths: (2 available, best #1, not advertised to EBGP peer)
  Advertised to non peer-group peers:
  10.3.0.5
  Local
    0.0.0.0 from 0.0.0.0 (10.3.0.6)
      Origin incomplete, metric 0, localpref 100, weight 32768,
valid,
      sourced, best
      Community: 387:31000 no-export
  Local
    10.3.0.2 (metric 128) from 10.3.0.5 (1.0.0.2)
      Origin incomplete, metric 0, localpref 100, valid, internal
      Originator: 1.0.0.2, Cluster list: 10.3.0.5
      Community: 387:31000 no-export
```

A) The primary link has come back, so the backup router now sees two alternate routes.

B) The primary link has not come back up, but the backup router still sees two alternate routes.

C) The first route is the route that the router itself has redistributed into BGP using the floating static route. This route is locally sourced by this AS and has been assigned a weight value of 32768.

D) The second route is the one received by IBGP from the primary edge router. This AS also sources this route, but no weight value is assigned.

Q6) What two things can you do to overcome the problems that occur when a floating static route is redistributed into BGP? (Choose two.)

A) You must raise the weight value.

B) You must lower the weight value.

C) You must set the AD at a higher value than all other routes.

D) You must assign local preference values, giving the floating static route a lower local preference value than the primary route.

 Copyright © 2003, Cisco Systems, Inc.

Q7) What are three characteristics of using static routes during load sharing of outgoing traffic? (Choose three.)

A) Outgoing traffic load sharing is easy to achieve.

B) Each customer router uses the closest customer edge router as the exit point.

C) Balanced load sharing is achieved if the customer edge routers are colocated.

D) Local preference values must be assigned, giving the floating static route a lower local preference value than the primary route.

Connecting a Multihomed Customer to a Single Service Provider

Quiz: Connecting a Multihomed Customer to a Single Service Provider

Q1) What are three responsibilities of the customer when the customer is multihomed to a single service provider? (Choose three.)

A) Customer edge routers must run IBGP between them.

B) The customer must advertise a default route.

C) The customer must conditionally advertise its assigned address space into BGP.

D) The customer edge routers must run EBGP with the provider.

Q2) Given the following router command output, what method has been used to influence return traffic in a primary/backup link implementation for this multihomed customer?

```
Provider# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.10.20.0/24    192.168.63.3          1000             0 100 100
i
*>                  192.168.64.4          2000             0 400 100
i
*> 30.30.30.0/24    192.168.63.3             0             0 100 I
*> 40.40.40.0/24    192.168.64.4             0             0 400 I
```

A) MED

B) local preference

C) weight

D) AS-path prepending

 Copyright © 2003, Cisco Systems, Inc.

Q3) What are three responsibilities of the provider router when supporting a multihomed customer? (Choose three.)

A) The provider must advertise a default route to the customer through BGP.

B) The provider must filter customer routes to verify that proper addressing is used.

C) The provider must remove the private AS number, if in use by the customer.

D) The provider must configure new AS-path filters to allow AS-path prepending; otherwise, a primary/backup link cannot be established.

Q4) What will occur if private AS numbers are advertised to the Internet?

A) The Internet will not be able to route packets.

B) Internet routers could drop routes based on BGP loop prevention mechanisms.

C) Customer load balancing will not function.

D) Customer configurations for the primary/backup link using AS-path prepending will not function.

Q5) What BGP configuration is required to properly implement a backup solution for a multihomed customer connected to a single provider? (Choose two.)

A) The customer should set local preference to influence outgoing route selection.

B) The customer should set the weight attribute to influence outgoing path selection.

C) The customer should set MED on each route to influence return path selection.

D) The customer should configure AS-path prepending to ensure proper outgoing path selection.

Copyright © 2003, Cisco Systems, Inc.

Q6) A customer router has been configured with maximum paths set to a value of 4. Given the following router command output, over how many links will the router need to perform load balancing?

```
router# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.10.20.0/24    192.168.63.3                           0 300 100
100 i
*>                  192.168.64.4                           0 400 100
i
*                   192.168.65.5                           0 500 100
i
*> 30.30.30.0/24    192.168.63.3             0             0 300 i
*> 40.40.40.0/24    192.168.64.4             0             0 400 i
```

A) The router will use only the path marked as "best" by BGP.

B) The router will perform load balancing over two paths to reach network 10.10.20.0/24.

C) The router will perform load balancing over three paths to reach network 10.10.20.0/24.

D) There is not enough information to determine the correct answer.

Q7) What three methods can you use to provide load sharing over network links between a multihomed customer and a single provider? (Choose three.)

A) advertising of split addressing space to the provider

B) configuring **ebgp-multihop** between the customer and the provider

C) use of the BGP **maximum-paths** command to perform load balancing over parallel links

D) configuring multiple static routes pointing to the provider

 Copyright © 2003, Cisco Systems, Inc.

Q8) Why is it not required to configure maximum paths under the BGP routing process when are performing load balancing with ebgp-multihop?

A) By default, BGP will perform load balancing over up to four paths, configurable up to six.

B) The static route or IGP process is responsible for load balancing in this configuration.

C) Configuring multihop enables maximum paths equal to the TTL setting of the **neighbor ebgp-multihop** command.

D) Configuring ebgp-multipath is a required component of ebgp-multihop load balancing.

Copyright © 2003, Cisco Systems, Inc.

Connecting a Multihomed Customer to Multiple Service Providers

Quiz: Connecting a Multihomed Customer to Multiple Service Providers

Q1) A multihomed customer is using AS number 65500 internally. The customer is connected to two different providers. Provider 1 (in AS 222) has assigned the customer an AS number of AS 65101. Provider 2 (in AS 333) has assigned the customer an AS number of AS 65201. Given that the customer will use AS number translation for its internal AS, what is the AS-path attribute, attached to routes originated in the customer network, when displayed on a router in the network of Provider 2?

A) 65550 i

B) 65201 i

C) 65201 65550 i

D) 333 65201 i

Q2) What three methods can you use to provide load sharing over network links between a multihomed customer and multiple providers? (Choose three.)

A) advertising of split addressing space to the provider

B) configuring of multiple static routes pointing to the provider

C) use of the BGP **maximum-paths** command to perform load sharing over parallel links

D) AS-path prepending to fine-tune the load-sharing configuration

Q3) What are three BGP configuration characteristics of a multihomed customer connected to multiple providers? (Choose three.)

A) The customer announces assigned addressing to its providers through BGP.

B) The customer announces a default route to its network through BGP.

C) The provider announces a default route, local routes, or full Internet routing to the customer via BGP.

D) The customer configures outbound filters to prevent its network from becoming a transit area.

 Copyright © 2003, Cisco Systems, Inc.

Q4) A multihomed customer is using AS number 1024 and is connected to two different providers (Provider 1: AS 222 and Provider 2: AS 333). The customer has configured MED to ensure a proper return path so that Provider 1 is the primary provider and Provider 2 is the backup provider. Unfortunately, return traffic continues to use the backup link. What is a possible cause of this problem?

A) The backup provider is ignoring the MED attribute on received routes.

B) The MED attribute cannot be sent to the backup provider because it is local to AS 1024 only.

C) The customer has not set the proper BGP communities to allow the primary and backup providers to correctly set the MED attribute.

D) MED cannot be used in this scenario, because it will not be advertised to providers upstream of Provider 2.

Q5) What are three important considerations for customers wishing to connect to multiple providers? (Choose three.)

A) The customer has to consider whether to use PA or PI address space.

B) The customer has to decide whether to use static routes or BGP to connect to upstream providers.

C) The customer has to decide whether to use a public AS number or a private AS number scheme.

D) The customer has to decide whether to perform load sharing or use a primary/backup implementation over redundant links.

Q6) Which AS number selection is the best possible choice for a customer multihomed to multiple providers?

A) a single public AS number

B) a single private AS number

C) two private AS numbers used in conjunction with AS number translation

D) multiple private AS numbers, one used internally by the customer and the others used in conjunction with AS number translation for each provider

Copyright © 2003, Cisco Systems, Inc.

Q7) Given the following router command output, what two methods have been configured to influence return traffic in a primary/backup link for this multihomed customer? (Choose two.)

```
Provider# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  10.10.20.0/24    192.168.63.3          2000             0 100 100
100 i
*>                  192.168.64.4                           0 300 100
i
*> 30.30.30.0/24    192.168.63.3             0             0 300 100
i
*> 40.40.40.0/24    192.168.64.4             0             0 100 i
```

A) MED

B) local preference

C) split address advertisement

D) AS-path prepending

 Copyright © 2003, Cisco Systems, Inc.

Module 5: BGP Transit Autonomous Systems

Transit Autonomous System Functions

Quiz: Transit Autonomous System Functions

Q1) Why is IBGP a mandatory component of a transit AS?

A) It is the only feasible way to ensure that all routers in the AS have consistent external routing information.

B) It eliminates the scalability issues of running an IGP within the transit AS.

C) Running IBGP on all routers is the only way to satisfy the filtering requirements of the transit AS.

D) An IGP is not capable of handling the potential routing loops in the transit AS.

Q2) How is EBGP used in a transit AS?

A) as a means of transporting customer routes across the transit backbone

B) to exchange routes between different autonomous systems and the transit AS

C) to enhance scalability by transporting IGP routes for the transit AS

D) as a means of injecting local routes into the transit backbone

Q3) Why is it not recommended to redistribute BGP routes into an IGP for use in a transit backbone?

A) Redistribution removes all BGP attributes needed to ensure optimal routing within the transit AS.

B) An IGP cannot enforce complex administrative policies and route selection rules.

C) IGPs cannot scale to the demands presented by the number of routes on the Internet.

D) IGPs are not stable when faced with a flapping network.

Copyright © 2003, Cisco Systems, Inc.

Q4) What are the two key functions of a transit AS? (Choose two.)

A) to filter out routes that do not belong to customers of the service provider

B) to provide Internet connectivity to customers of the service provider

C) to propagate routes between remote autonomous systems

D) to route packets between remote networks

Q5) How are BGP routes sent across the transit backbone?

A) by redistributing BGP into an IGP and then back into BGP

B) through the use of IBGP

C) by establishing EBGP sessions between all routers in the transit backbone

D) by redistributing connected routes at the edge of the transit backbone

 Copyright © 2003, Cisco Systems, Inc.

IBGP and EBGP Interaction in a Transit Autonomous System

Quiz: IBGP and EBGP Interaction in a Transit Autonomous System

Q1) What two statements are true regarding the AS-path attribute as it relates to IBGP? (Choose two.)

A) Each router in the AS appends its AS number to the AS path on outgoing BGP updates.

B) The AS path inside an AS will be empty for routes originating inside a neighboring AS.

C) The AS-path attribute is not used to detect routing loops inside an AS.

D) The AS-path attribute is not modified within the AS.

Q2) Why is it recommended that loopback interfaces be used to form IBGP neighbor sessions?

A) Using loopback interfaces reduces router memory resource requirements.

B) Using loopback interfaces reduces router CPU resource requirements.

C) Using loopback interfaces ensures IBGP session stability.

D) Using loopback interfaces is more secure than using the physical interface.

Q3) How is the BGP next-hop attribute processed over an IBGP connection?

A) The next-hop address is set to the address of the receiving router.

B) The next-hop address is not modified over the IBGP session.

C) The next-hop address is set to the IP address of the nearest EBGP peer.

D) The next-hop attribute is set to the IP address of the nearest EBGP peer; if no external AS connection has been configured, the next hop is set to the default gateway configured on the router.

Copyright © 2003, Cisco Systems, Inc.

Q4) Which two statements are true of the full-mesh requirement in IBGP? (Choose two.)

A) The IBGP mesh requires a logical full mesh.

B) A physical full mesh must be maintained within the IBGP AS.

C) Due to BGP split horizon, no router can relay IBGP information within the AS.

D) All routers within the AS must be directly connected to ensure correct delivery of BGP routing information.

Q5) What three statements are true regarding the next-hop-self configuration in BGP? (Choose three.)

A) Changing the next-hop attribute might cause suboptimal routing.

B) The configuration changes how the next-hop attribute is processed at edge routers.

C) The configuration announces the local IP address as the BGP next hop in outgoing updates sent to the specified neighbor.

D) The configuration removes the requirement for the IGP to carry reachability information for intra-AS destinations.

Q6) What are three differences between IBGP and EBGP sessions? (Choose three.)

A) Route selection rules slightly prefer IBGP routes.

B) Routes learned from IBGP peer are not advertised to other IBGP peers.

C) EBGP peers are directly connected, and IBGP peers are usually distant.

D) By default, no BGP attributes are changed in IBGP updates.

 Copyright © 2003, Cisco Systems, Inc.

Packet Forwarding in Transit Autonomous Systems

Quiz: Packet Forwarding in Transit Autonomous Systems

Q1) What are two reasons why you must run IBGP on all routers within a transit backbone? (Choose two.)

A) so routers can properly forward packets toward all external destinations

B) to ensure that a full mesh exists between all routers in the AS

C) to allow routers to properly process the BGP next-hop attribute

D) because IGPs cannot scale large enough to handle redistribution of BGP routes

Q2) If a transit backbone has IBGP running on all routers, what are two reasons why it is still necessary to use an IGP? (Choose two.)

A) to provide routing information needed to establish the IBGP sessions

B) to resolve next-hop references used in recursive routing

C) so that BGP routes can be properly transported through the AS

D) to provide user workstations with a network default gateway

Q3) What is the AD of the following protocols? (Fill in the blanks.)

A) IBGP __200_____

B) EBGP __20_____

C) OSPF __110_____

D) IS-IS __115_____

E) RIP __120_____

Q4) What are two reasons why the AD is an important consideration for BGP network design? (Choose two.)

A) The AD affects how routes are selected for use in the IP routing table.

B) The AD controls how routing information is entered into the BGP table.

C) If a route is advertised by both an IGP and through EBGP, the router will prefer the external route.

D) AD is not a large concern to BGP design, because the router will always choose the route advertised by the protocol best suited to reach the destination.

Q5) With regard to recursive route lookups, what are two ways in which CEF is different from traditional Cisco IOS switching mechanisms such as route caching? (Choose two.)

A) Traditional Cisco IOS switching mechanisms wait for the first packet to arrive before recursive lookup can take place.

B) New entries in the IP routing table will trigger a recursive lookup in traditional Cisco IOS switching mechanisms.

C) CEF prebuilds a complete IP forwarding table based on the IP routing table.

D) CEF will build a FIB directly from the entries in the BGP table prior to any BGP packets arriving at the router.

 Copyright © 2003, Cisco Systems, Inc.

Configuring a Transit Autonomous System

Quiz: Configuring a Transit Autonomous System

Q1) When you are configuring the BGP neighbor session, what differentiates an EBGP neighbor from an IBGP neighbor?

A) The keyword **internal** at the end of the **neighbor** command.

B) IBGP neighbors will have the same AS number specified.

C) A description for the neighbor must be attached with the **neighbor description** command.

D) Directly connected neighbors will automatically form an EBGP session.

Q2) What two steps are required to use a loopback interface for IBGP peering sessions? (Choose two.)

A) Ensure that the loopback interfaces are reachable through an IGP.

B) Ensure that the two neighbors must be directly attached.

C) Verify that each router has multiple physically redundant paths.

D) Configure a **neighbor** statement with the **update-source** command.

Q3) Why is it important to disable BGP synchronization in a transit backbone?

A) IGPs can support the routing requirements of full Internet routing, and hence synchronization is no longer necessary.

B) Because BGP redistribution into an IGP is no longer practical, enabling the synchronization feature is no longer applicable.

C) Synchronization requires all BGP transit routes to be explicitly mapped to an exit point, creating too much administrative overhead.

D) Synchronization requires BGP attributes to be properly mapped to IGP metrics in order for BGP routing across the transit backbone to function properly, creating too much overhead.

Copyright © 2003, Cisco Systems, Inc.

Q4) What are two negative ramifications of the full-mesh requirement imposed by IBGP? (Choose two.)

A) administratively difficult to apply an AS-wide routing policy

B) requires the use of next-hop-self for proper routing to external destinations

C) large number of TCP sessions

D) unnecessary duplication of routing traffic

Q5) What are two scalability tools that you can use to overcome the full-mesh requirement for IBGP sessions? (Choose two.)

A) confederations

B) floating static routes

C) route reflectors

D) disabling BGP synchronization

 Copyright © 2003, Cisco Systems, Inc.

Module 6: Scaling Service Provider Networks

Scaling IGP and BGP in Service Provider Networks

Quiz: Scaling IGP and BGP in Service Provider Networks

Q1) What three characteristics are common to typical service provider networks? (Choose three.)

A) The provider network uses two IGPs, one for customer routes and one for internal provider routes.

B) Service providers exchange routes with other providers using BGP.

C) Service providers run IBGP within their network in addition to their IGP requirements.

D) Service providers typically use either static routes or EBGP with their customers.

Q2) What is the typical role of an IGP within a service provider network?

A) The IGP carries customer routes for redistribution into BGP at the provider edge.

B) The IGP advertises a default route to customers of the service provider.

C) The IGP resolves next-hop IP addresses.

D) The IGP carries BGP routes across the provider network.

Q3) Why should you avoid the use of private IP addressing in service provider networks?

A) Private addressing can prevent customer network troubleshooting utilities such as traceroute from functioning correctly.

B) Private IP addressing is not allowed on the Internet and will not function in a service provider network.

C) Private IP addressing prevents the service provider from properly summarizing customer routes if they are also using private address space.

D) Private IP addressing prevents service provider applications such as MPLS from operating properly in an Internet-supporting environment.

Copyright © 2003, Cisco Systems, Inc.

Q4) What three are key to properly scaling BGP in a service provider environment? (Choose three.)

A) IBGP full-mesh scaling tools to reduce duplicate traffic within the AS

B) summarization of customer routes to reduce the number of prefixes carried

C) improving BGP convergence time by using the IGP for route propagation within the provider AS

D) proper scaling of the AS-wide routing policy to ease administration and maintenance requirements

 Copyright © 2003, Cisco Systems, Inc.

Introduction to Route Reflectors

Quiz: Introduction to Route Reflectors

Q1) What is the main problem solved by implementing BGP route reflectors?

A) the large number of routes carried in the IGP when BGP is deployed

B) the ability for BGP to scale a single AS in a large network

C) the need for a homogeneous method of applying policies to routes carried through an AS

D) the lack of a service level agreement supporting features and the need to deploy these features with greater ease

Q2) How does a route reflector modify the IBGP split-horizon rule?

A) Route reflectors forward EBGP updates onto all peers (IBGP and EBGP).

B) Route reflectors treat all neighbors as EBGP peers, eliminating the IBGP mesh requirements.

C) Route reflectors forward IBGP updates from clients to other IBGP neighbors.

D) Route reflectors append the cluster-ID to the AS path, allowing peers to be treated as EBGP neighbors.

Q3) Why are redundant route reflectors mandatory in any high-availability network design?

A) because all neighbors peer with the route reflector, and a the large number of neighbors can make the route reflector router unstable

B) because EBGP peers can inject BGP updates into the AS only through the route reflector

C) because route reflectors maintain more routing information, making them more prone to congestion and failure

D) because clients can form IBGP relationships only with the route reflector

Copyright © 2003, Cisco Systems, Inc.

Q4) What is the main reason for implementing redundant route reflectors with clusters?

A) to eliminate routing loops in redundant configurations

B) to limit the number of neighbor sessions with each route reflector

C) to provide another scalability mechanism targeted at removing the IBGP full-mesh requirement

D) to enhance security within the AS

Q5) How does the originator-ID attribute assist in the elimination of routing loops caused by redundant route reflector designs?

A) If the originator-ID matches the router-ID of the reflector, local preference is set on the route to make it a backup.

B) The originator-ID attribute is set to the cluster-ID to ensure that a route traverses the AS only one time.

C) A router receiving a route with the originator-ID matching its router-ID will ignore that route.

D) The originator-ID allows the router to know if the route originated locally or from an external source so that administrative distance rules for the route can be verified.

 Copyright © 2003, Cisco Systems, Inc.

Network Design with Route Reflectors

Quiz: Network Design with Route Reflectors

Q1) What can occur if a client has IBGP neighbor relationships with other routers not configured as route reflectors?

A) This is an invalid configuration.

B) The client will notify the route reflector and be promoted to a route reflector as well.

C) Routing black holes can occur, causing lost traffic inside the AS.

D) Unnecessary routing traffic will be generated.

Q2) What potential problem can occur if a client does not have an IBGP session with all route reflectors in a cluster?

A) This is an invalid configuration.

B) The client might not receive all BGP routes.

C) EBGP routes received by the client will not be distributed properly throughout the AS.

D) Duplicate routing traffic will be sent to the client.

Q3) What problem are hierarchical route reflectors designed to solve?

A) lack of a consistent application of security and routing policies throughout the AS

B) scalability of autonomous systems in very large routing domains

C) routing loops caused by redundant cluster configurations

D) administrative overhead when you are implementing router reflector network designs

Copyright © 2003, Cisco Systems, Inc.

Introduction to Confederations

Quiz: Introduction to Confederations

Q1) What is the main problem solved by implementing BGP confederations?

A) the large number of routes carried in the IGP when BGP is deployed

B) the ability for BGP to scale a single AS in a large network

C) the need for a homogeneous method of applying policies to routes carried through an AS

D) the lack of a service level agreement supporting features and the need to deploy these features with greater ease

Q2) Although confederations eliminate the need for a fully meshed topology within the AS, where does the BGP full-mesh requirement still apply?

A) To all EBGP neighbor sessions.

B) Inside each member AS.

C) Between the member autonomous systems contained in the confederation.

D) The IBGP full-mesh requirement no longer applies when confederations are used in an AS.

Q3) How does an IBGP router receiving the AS-path attribute in a BGP update determine if the route has crossed a member-AS within a confederation?

A) By the presence of the confederation bit in the flag field of the BGP update.

B) Because the AS-path attribute will contain only the AS number of the ingress EBGP peer.

C) The member-AS numbers will be indicated by the presence of parentheses surrounding the AS number entry.

D) The IBGP router cannot determine whether the route has crossed a member-AS because the AS number of each AS boundary crossed is appended to the AS-path attribute.

 Copyright © 2003, Cisco Systems, Inc.

Q4) How does an EBGP router receiving the AS-path attribute in a BGP update determine if the route has crossed a member-AS within a confederation?

A) By the presence of the confederation bit in the flag field of the BGP update.

B) Because the AS-path attribute will contain only the AS number of the ingress EBGP peer.

C) The member-AS numbers will be indicated by the presence of parentheses surrounding the AS number entry.

D) The EBGP router cannot determine whether the route has crossed a member-AS because the member-AS entries are removed from the AS path prior to exiting the confederation.

Q5) Why is it not possible for a router that does not support BGP confederations to operate inside an AS configured as a confederation?

A) The router will believe that the AS path is longer than the actual AS path and route incorrectly.

B) The router will be unable to interpret the AS-path attribute and terminate its BGP session with that peer.

C) The router will automatically convert the intraconfederation AS numbers to the external AS number of the confederation, causing an AS number mismatch.

D) The router will process BGP updates as normal because it has to be aware only of the member-AS to which it belongs, causing incorrect routing information to propagate through the AS.

Q6) Which three IBGP properties are retained within the confederation even though EBGP sessions between member autonomous systems are formed? (Choose three.)

A) local preference

B) MED

C) weight

D) next-hop

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

Module 7: Optimizing BGP Scalability

Improving BGP Convergence

Quiz: Improving BGP Convergence

Q1) What are three characteristics of a converged BGP network? (Choose three.)

A) The input queue and output queue for all peers is 0.

B) All routes in the BGP table have been installed in the routing table.

C) The table version for all peers equals the table version of the BGP table.

D) All routes have been accepted.

Q2) Which two of the following modifications result in improved BGP convergence? (Choose two.)

A) increasing the default value of BGP hold time

B) lowering the default value of BGP scan time

C) increasing the default value of the neighbor advertisement intervals

D) lowering the default value of the neighbor advertisement intervals

Q3) What is the main task of the BGP scanner process?

A) sends routing updates to BGP neighbors

B) walks the BGP table for routes to enter into the IP routing table

C) confirms the reachability of BGP next hops

D) scans the router configuration to establish and maintain BGP neighbors

Copyright © 2003, Cisco Systems, Inc.

Q4) One of your BGP core routers is experiencing periodic slow responses to ping packets directed to it from the network management console. The router has just been configured to receive full Internet routes, and you suspect that the BGP router process is causing CPU utilization issues in the core router. Which two router commands should you use to confirm your suspicion? (Choose two.)

A) **show ip route**

B) **show ip bgp summary**

C) **show process cpu**

D) **show memory**

Q5) The output of a **show interfaces fastethernet 0/0** command follows:

```
Fast Ethernet0 is up, line protocol is up
  Hardware is DEC21140, address is 0000.0c0c.1111 (bia
0002.eaa3.5a60)
  Internet address is 112.64.101.17 255.255.255.240
  MTU 1460 bytes, BW 100000 Kbit, DLY 100 usec, rely 255/255, load
200/255
  Encapsulation ARPA, loopback not set, keepalive not set, hdx,
100BaseTX
  ARP type: ARPA, ARP Timeout 4:00:00
  Last input never, output 0:00:16, output hang 0:28:01
  Last clearing of "show interface" counters 0:20:05
  Output queue 25/40, 0 drops; input queue 50/500, 1470 drops
  5 minute input rate 21666400 bits/sec, 1855 packets/sec
  5 minute output rate 72221 bits/sec, 618 packets/sec
```

How has the interface been modified to improve BGP convergence?

A) The output queue has been decreased to expedite packet forwarding out the fast Ethernet interface.

B) The drop threshold of the input queue has been set to begin randomly discarding packets after the queue reaches 50 packets deep.

C) PMTU discovery has been enabled, setting the interface MSS to 1460 bytes.

D) The size of the input queue has been increased to support up to 500 incoming packets.

 Copyright © 2003, Cisco Systems, Inc.

Q6) Refer to the following Cisco IOS router output:

```
router# show ip bgp summary
BGP router identifier 172.16.0.4, local AS number 1
BGP table version is 16, main routing table version 16
20 network entries and 20 paths using 2826 bytes of memory
8 BGP path attribute entries using 480 bytes of memory
7 BGP AS-PATH entries using 168 bytes of memory
3 BGP community entries using 72 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
3 BGP filter-list cache entries using 36 bytes of memory
BGP activity 20/0 prefixes, 24/4 paths, scan interval 120 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
172.16.0.1      4     1      30      30       16    0    0 00:23:13
5
172.16.0.2      4     1      33      30       16    0    0 00:23:15
5
172.16.0.3      4     1      27      30       16    0    0 00:23:14
5
192.168.21.99   4    99      31      35       16    0    0 00:23:04
5
```

What two parameters would indicate that the BGP network has converged? (Choose two.)

A) The TblVer for all neighbors is 16.

B) V is set to 4 for all neighbors.

C) The InQ and OutQ for all neighbors is 0.

D) All neighbors are in the Established state and have the same PfxRcd value.

Q7) Using the command output from question 6, identify how frequently the BGP scanner process will run on the router.

A) By default, the process will run every 60 seconds.

B) The process has run 16 times and will run again when the next BGP update arrives.

C) The process will run on this router every 120 seconds.

D) It cannot be determined from this output.

Q8) What are two potential issues caused by modifying the default scan time and advertisement interval on a BGP router? (Choose two.)

A) Router CPU resources can be exhausted.

B) Router memory resources can be depleted.

C) Routing loops are more likely.

D) BGP could converge faster than the IGP, causing network black holes.

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise Solutions

The laboratory exercise solutions are listed here.

Module 1: BGP Overview

Laboratory Exercise 1-1: Configuring BGP

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

```
WGxR1#configure terminal
WGxR1(config)#router bgp x
WGxR1(config-router)#neighbor 192.168.20.20 remote-as 20
WGxR1(config-router)#network 192.168.x.0
WGxR1(config-router)#network 197.x.0.0 mask 255.255.0.0
WGxR1(config-router)#exit
WGxR1(config)#ip route 192.168.x.0 255.255.255.0 null 0 250
WGxR1(config)#ip route 197.x.0.0 255.255.0.0 null 0 250
WGxR1(config)#end
```

Also, make sure that you announce a default route into your network. If you are running OSPF in your network, the command sequence to use is:

```
WGxR1#configure terminal
WGxR1(config)#router ospf process-id
WGxR1(config-router)#default-information originate always
```

Copyright © 2003, Cisco Systems, Inc.

Answers to Review Questions:

Q1) What do you need in order to be able to propagate classful networks?

There must be at least one subnet in the routing table, and the network listed in the BGP process with no mask attached to it.

Q2) What do you need to be able to propagate classless networks (supernets or subnets)?

The prefix must be listed in the BGP process together with its proper mask. The exact corresponding entry must also be present in the IP routing table.

Q3) Why do some networks, received from router "Good," have a next-hop address pointing to other routers?

Router "Good" performs next-hop processing and sets the next hop to router "Cheap" for routes received from "Cheap" and propagated to you.

Q4) What command would you use to see if a neighbor is sending you any updates and how many?

You would use the **show ip bgp neighbor** command.

Laboratory Exercise 1-2: Configuring Route Redistribution in BGP

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

```
WGxR1#configure terminal
WGxR1(config)#router bgp x
WGxR1(config-router)#no network 192.168.x.0
WGxR1(config-router)#no network 197.x.0.0 mask 255.255.0.0
WGxR1(config-router)#redistribute <igp> <pid> route-map SetOrigin
WGxR1(config)#access-list 1 deny   192.168.20.0
WGxR1(config)#access-list 1 deny   192.168.21.0
WGxR1(config)#access-list 1 permit any
WGxR1(config)#route-map SetOrigin permit 10
WGxR1(config-route-map)#match ip address 1
WGxR1(config-route-map)#set origin igp
WGxR1(config-route-map)#end
WGxR1#
```

 Copyright © 2003, Cisco Systems, Inc.

Answers to Review Questions:

Q1) What is the major difference between this implementation and the previous one? Which is better and why?

The previous implementation is better for service provider environments because it is more controlled. This implementation is better for enterprise networks because it is more dynamic.

Q2) What precautions do you have to take when using redistribution?

You should always filter redistributed routes with a route-map to make sure that unwanted networks are not announced into BGP.

Laboratory Exercise 1-3: Configuring BGP Aggregation

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

```
WGxR1(config)#router bgp x
WGxR1(config-router)#aggregate-address 197.x.0.0 255.255.0.0 summary-
only
WGxR1(config-router)#aggregate-address 197.x.0.0 255.255.248.0
summary-only
WGxR1(config-router)#aggregate-address 197.x.8.0 255.255.252.0
summary-only
WGxR1(config-router)#end
WGxR1#
```

Answers to Review Questions:

Q1) Do you see all your prefixes on the provider router? Why?

You should see only the configured aggregates and the routes not within the aggregation range on the provider router. All the other routes should be suppressed.

Q2) What do you need to be able to generate and propagate aggregates?

The aggregate address has to be configured in the BGP process, and a prefix within the aggregate range has to be present in the BGP table.

Before Aggregation

```
> 197.1.1.0        192.168.20.1             0             0 1 i
> 197.1.2.0        192.168.20.1             0             0 1 i
> 197.1.3.0        192.168.20.1             0             0 1 i
> 197.1.4.0        192.168.20.1             0             0 1 i
> 197.1.5.0        192.168.20.1             0             0 1 i
> 197.1.6.0        192.168.20.1             0             0 1 i
> 197.1.7.0        192.168.20.1             0             0 1 i
> 197.1.8.0        192.168.20.1             0             0 1 I
```

After Aggregation

```
*> 197.1.0.0/21     192.168.20.1                           0 1 i
*> 197.1.0.0/16     192.168.20.1                           0 1 i
*> 197.1.8.0/22     192.168.20.1                           0 1 I
```

 Copyright © 2003, Cisco Systems, Inc.

Module 2: Route Selection Using Policy Controls

Laboratory Exercise 2-1: Multihomed BGP Networks

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

```
WGxR1(config)#router bgp x
WGxR1(config-router)#neighbor 192.168.20.22 remote-as 22
WGxR1(config-router)#neighbor 192.168.20.22 weight 100
WGxR1(config-router)#end
WGxR1#
```

Answers to Review Questions:

Q1) What can happen if a multihomed AS is passing routing information, learned from one neighbor, to another neighbor?

The multihomed AS passing routes between BGP neighbors can become a transit AS.

Q2) Why do some prefixes have two paths but both use the same next-hop address?

Both service providers are connected to the same subnet as all the customers, and the BGP next-hop processing optimizes the data flow.

Laboratory Exercise 2-2: AS-Path Filters

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

```
WGxR1(config)#ip as-path access-list 1 deny _214$
WGxR1(config)#ip as-path access-list 1 permit .*
WGxR1(config)#ip as-path access-list 2 deny _213$
WGxR1(config)#ip as-path access-list 2 permit .*
WGxR1(config)#ip as-path access-list 3 permit ^$
WGxR1(config)#router bgp x
WGxR1(config-router)#neighbor 192.168.20.20 filter-list 1 in
WGxR1(config-router)#neighbor 192.168.20.22 filter-list 2 in
WGxR1(config-router)#neighbor 192.168.20.20 filter-list 3 out
WGxR1(config-router)#neighbor 192.168.20.22 filter-list 3 out
WGxR1(config-router)#end
WGxR1#clear ip bgp *
```

Copyright © 2003, Cisco Systems, Inc.

Answers to Review Questions:

Q1) By this time, a the BGP table should contain a large number of prefixes. What regular expression would you use with the **show ip bgp command** on WGxR1 to view prefixes originated by your AS?

Use **show ip bgp regexp ^$**.

```
wg1r1#sh ip bgp regexp ^$
BGP table version is 44, local router ID is 197.1.8.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
inter
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric LocPrf Weight Path
*> 192.168.1.0      0.0.0.0                  0         32768 i
*> 197.1.0.0/21     0.0.0.0                            32768 i
*> 197.1.0.0/16     0.0.0.0                            32768 i
s> 197.1.1.0        0.0.0.0                  0         32768 i
s> 197.1.2.0        0.0.0.0                  0         32768 i
s> 197.1.3.0        0.0.0.0                  0         32768 i
s> 197.1.4.0        0.0.0.0                  0         32768 i
s> 197.1.5.0        0.0.0.0                  0         32768 i
s> 197.1.6.0        0.0.0.0                  0         32768 i
s> 197.1.7.0        0.0.0.0                  0         32768 i
s> 197.1.8.0        0.0.0.0                  0         32768 i
*> 197.1.8.0/22     0.0.0.0                            32768 I
```

Q2) What regular expression would you use on router “Good” or “Cheap” to view prefixes originating in your AS?

Use **show ip bgp regexp _1$.**

AS is 1.

```
Good# sh ip bgp regexp _1$
BGP table version is 111, local router ID is 199.199.199.199
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric LocPrf Weight Path
*  192.168.1.0      192.168.20.1                           0 22 1 i
*>                  192.168.20.1             0             0 1 i
*  197.1.0.0/21     192.168.20.1                           0 22 1 i
*>                  192.168.20.1                           0 1 i
*  197.1.0.0/16     192.168.20.1                           0 22 1 i
*>                  192.168.20.1                           0 1 i
*  197.1.8.0/22     192.168.20.1                           0 22 1 i
*>                  192.168.20.1                           0 1 i
```

 Copyright © 2003, Cisco Systems, Inc.

Q3) What regular expression would you use on router "Good" or "Cheap" to view prefixes received from your AS?

Use **show ip bgp regexp ^x_.**

```
Good#sh ip bgp regexp ^1_
BGP table version is 111, local router ID is 199.199.199.199
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric LocPrf Weight Path
*> 192.168.1.0      192.168.20.1             0             0 1 i
*> 197.1.0.0/21     192.168.20.1                           0 1 i
*> 197.1.0.0/16     192.168.20.1                           0 1 i
*> 197.1.8.0/22     192.168.20.1                           0 1 i
```

Q4) How can you test your AS-path filters before applying them and clearing BGP neighbors?

You can test the AS-path filters with the **show ip bgp filter-list** *filter* command.

Laboratory Exercise 2-3: Prefix-List Filters

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

```
WGxR1(config)#ip prefix-list FromGood deny 192.0.0.0/3 ge 24
WGxR1(config)#ip prefix-list FromGood deny 192.168.0.0/16 le 32
WGxR1(config)#ip prefix-list FromGood deny 172.16.0.0/12 le 32
WGxR1(config)#ip prefix-list FromGood deny 10.0.0.0/8 le 32
WGxR1(config)#ip prefix-list FromGood permit 0.0.0.0/0 le 32
WGxR1(config)#ip prefix-list FromCheap deny 128.0.0.0/2 ge 17
WGxR1(config)#ip prefix-list FromCheap deny 192.168.0.0/16 le 32
WGxR1(config)#ip prefix-list FromCheap deny 172.16.0.0/12 le 32
WGxR1(config)#ip prefix-list FromCheap deny 10.0.0.0/8 le 32
WGxR1(config)#ip prefix-list FromCheap permit 0.0.0.0/0 le 32
WGxR1(config)#router bgp x
WGxR1(config-router)#neighbor 192.168.20.20 prefix-list FromGood in
WGxR1(config-router)#neighbor 192.168.20.22 prefix-list FromCheap in
WGxR1(config-router)#end
WGxR1#clear ip bgp *
```

Answers to Review Questions:

Q1) Are neighboring workgroups still reachable?

Yes, They are still reachable.

Q2) Why does router WGxR1 still accept class B networks from router "Cheap"?

Only subnets of class B networks are filtered, not the class B networks themselves.

Laboratory Exercise 2-4: Implementing Changes in BGP Policy

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

```
wg1r1#sh ip bgp ne 192.168.20.20 received
% Inbound soft reconfiguration not enabled

wg1r1#sh ip bgp ne 192.168.20.22 received
% Inbound soft reconfiguration not enabled

wg1r1(config)#router bgp 1
wg1r1(config-router)#neighbor 192.168.20.20 soft-reconfiguration
inbound
wg1r1(config-router)#neighbor 192.168.20.22 soft-reconfiguration
inbound
wg1r1(config-router)#exit
wg1r1(config)#exit
```

Answers to Review Questions:

Q1) What command do you use to show which entries in the BGP table of the local router have been propagated to a specific neighbor?

show ip bgp neighbors *ip-address* **advertised**

Q2) What command do you use to display which of the routes in the local BGP table have been received and accepted from an indicated neighbor?

show ip bgp neighbors *ip-address* **routes**

 Copyright © 2003, Cisco Systems, Inc.

Module 3: Route Selection Using Attributes

Laboratory Exercise 3-1: Influencing BGP Route Selection with Weights

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

```
WGxR1(config)#ip as-path access-list 10 permit _37_|_213_
WGxR1(config)#route-map SetWeight permit 10
WGxR1(config-route-map)#match as-path 10
WGxR1(config-route-map)#set weight 300
WGxR1(config-route-map)#exit
WGxR1(config)#route-map SetWeight permit 1000
WGxR1(config-route-map)#exit
WGxR1(config)#router bgp x
WGxR1(config-router)#neighbor 192.168.20.20 weight 100
WGxR1(config-router)#neighbor 192.168.20.22 weight 200
WGxR1(config-router)#neighbor 192.168.20.20 route-map SetWeight in
WGxR1(config-router)#end
WGxR1#clear ip bgp *
WGxR1#
```

Answers to Review Questions:

Q1) Did all paths automatically get a weight of 100 or 200? Why not? What did you have to do?

The BGP sessions must be reset in order to apply newly configured weights to incoming BGP routes.

Q2) Name some parameters and attributes used for best-path selection.

The following BGP parameters and attributes are used for BGP best-path selection in the order specified: weight, local preference, AS-path length, origin code, multi-exit discriminator (MED).

Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise 3-2: BGP Local Preference

Use the following command sequence on your workgroup routers as indicated (replace x with your workgroup number):

Configuration of router WGxR2:

```
WGxR2(config)#route-map SetLP
WGxR2(config-route-map)#set local-preference 200
WGxR2(config-route-map)#exit
WGxR2(config)#router bgp 1
WGxR2(config)#router# no synchronization
WGxR2(config)# neighbor 192.168.1.1 remote-as 1
WGxR2(config-router)#network 192.168.x.0
WGxR2(config-router)#neighbor 192.168.3x.2 remote-as 20
WGxR2(config-router)#neighbor 192.168.3x.2 route-map SetLP in
WGxR2(config-router)#exit
WGxR2(config)#int s0/0.200 point-to-point
WGxR2(config-subif)#ip address 192.168.31.1 255.255.255.252
WGxR2(config-subif)#frame-relay interface-dlci 201
WGxR2(config-fr-dlci)#exit
WGxR2(config-subif)#exit
WGxR2(config)#exit
```

Configuration of router WGxR2:

```
WGxR1(config)router bgp 1
WGxR1(config-router)#no synchronization
WGxR1(config-router)#neighbor 192.168.1.2 remote-as 1
WGxR1(config-router)#no neighbor 192.168.20.20 weight 100
WGxR1(config-router)#no neighbor 192.168.20.22 weight 200
WGxR1(config-router)#no neighbor 192.168.20.20 route-map SetWeight in
WGxR1(config-router)#end
```

Depending on your IGP setup, you might need to perform additional configuration of your IGP to insert 192.168.3x.0/30 into your IGP:

```
WGxR2(config)#router ospf 1
WGxR2(config-router)#network 192.168.3x.1 0.0.0.0 area 0
WGxR2(config-router)#passive-interface serial 0/0.200
```

Answers to Review Questions

Q1) Is routing between router WGxR1 and “Good” symmetrical?

Initially, the routing is not symmetrical, because the router “Good” prefers more stable EBGP routes received from WGxR1 and the router WGxR1 prefers routes received from WGxR2.

Q2) Which routers receive the local preference attribute?

Only routers within the AS receive the local preference attribute.

 Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise 3-3: BGP Multi-Exit Discriminator

Use the following command sequence on your router WGxR1 (replace x with your workgroup number):

Configuration of router WGxR1:

```
WGxR1(config)#route-map SetMED
WGxR1(config-route-map)#set metric 150
WGxR1(config-route-map)#exit
WGxR1(config)#router bgp 1
WGxR1(config-router)#neighbor 192.168.20.20 route-map SetMED out
WGxR1(config-router)#end
WGxR1#
WGxR1#clear ip bgp *
```

Configuration of router WGxR2:

```
WGxR2(config)#route-map SetMED
WGxR2(config-route-map)#set metric 50
WGxR2(config-route-map)#exit
WGxR2(config)#router bgp 1
WGxR2(config-router)#neighbor 192.168.3x.2 route-map SetMED out
WGxR2(config-router)#end
WGxR2#
WGxR2#clear ip bgp *
```

Answers to Review Questions

Q1) Which parameters and attributes have to be equal before MED is compared to select the best path?

MED is used only to compare routes received from the same AS that have equal weight, local preference, AS-path length, and origin code.

Q2) What is the default value of MED?

For networks originated in the local AS, the default MED value is the IGP metric from the IP routing table. For transit networks, MED is removed by the router in outgoing EBGP updates.

Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise 3-4: BGP Communities

Use the following command sequence on your router WGxR2 (replace x with your workgroup number):

Configuration of router WGxR2:

```
WG1R2(config)#route-map SetCommunity
WG1R2(config-route-map)#set community 1:20
WG1R2(config-route-map)#exit
WG1R2(config)#router bgp 1
WG1R2(config-router)#neighbor 192.168.31.2 route-map SetCommunity out
WG1R2(config-router)#neighbor 192.168.31.2 send-community
WG1R(config-router)#exit
WG1R2(config)#ip bgp-community new-format
WG1R2(config)#exit
 WG1R2#clear ip bgp *
```

Answers to Review Questions

Q1) What do you have to do to enable community propagation?

Enable BGP community propagation with the **neighbor send-community** command.

Q2) What mechanisms can you use to match or set communities?

Only route-maps allow matching or setting of BGP communities.

 Copyright © 2003, Cisco Systems, Inc.

Module 5: BGP Transit Autonomous Systems

Laboratory Exercise 5-1: Configure the BGP Transit Autonomous System

Use the following command sequence on your router WGxR1:

```
WGxR1(config)#router bgp x
WGxR1(config-router)#no synchronization
WGxR1(config-router)#neighbor 192.168.20.20 remote-as 20
WGxR1(config-router)#neighbor 192.168.20.22 remote-as 22
WGxR1(config-router)#neighbor 197.x.2.1 remote-as x
WGxR1(config-router)#neighbor 197.x.2.1 update-source loopback 0
WGxR1(config-router)#neighbor 197.x.4.1 remote-as x
WGxR1(config-router)#neighbor 197.x.4.1 update-source loopback 0
WGxR1(config-router)#neighbor 197.x.6.1 remote-as x
WGxR1(config-router)#neighbor 197.x.6.1 update-source Loopback 0
WGxR1(config-router)#network 192.168.1.0
WGxR1(config-router)#network 197.1.0.0 mask 255.255.0.0
WGxR1(config-router)#end
```

Use the following command sequence on your router WGxR2:

```
WGxR2(config)#router bgp x
WGxR2(config-router)#no synchronization
WGxR2(config-router)#neighbor 197.x.1.1 remote-as x
WGxR2(config-router)#neighbor 197.x.1.1 update-source Loopback0
WGxR2(config-router)#neighbor 197.x.4.1 remote-as x
WGxR2(config-router)#neighbor 197.x.4.1 update-source Loopback0
WGxR2(config-router)#neighbor 197.x.6.1 remote-as x
WGxR2(config-router)#neighbor 197.x.6.1 update-source Loopback0
WGxR2(config-router)#end
```

Use the following command sequence on your router WGxR3:

```
WGxR3(config)#router bgp x
WGxR3(config-router)#no synchronization
WGxR3(config-router)#neighbor 197.x.1.1 remote-as x
WGxR3(config-router)#neighbor 197.x.1.1 update-source Loopback0
WGxR3(config-router)#neighbor 197.x.2.1 remote-as x
WGxR3(config-router)#neighbor 197.x.2.1 update-source Loopback0
WGxR3(config-router)#neighbor 197.x.6.1 remote-as x
WGxR3(config-router)#neighbor 197.x.6.1 update-source Loopback0
WGxR3(config-router)#end
```

Use the following command sequence on your router WGxR4:

```
WGxR4(config)#router bgp x
WGxR4(config-router)#no synchronization
WGxR4(config-router)#neighbor 192.168.21.99 remote-as 99
WGxR4(config-router)#network 192.168.1.0
WGxR4(config-router)#network 197.1.0.0 mask 255.255.0.0
```

Copyright © 2003, Cisco Systems, Inc.

```
WGxR4(config-router)#neighbor 197.x.1.1 remote-as x
WGxR4(config-router)#neighbor 197.x.1.1 update-source Loopback 0
WGxR4(config-router)#neighbor 197.x.2.1 remote-as x
WGxR4(config-router)#neighbor 197.x.2.1 update-source loopback 0
WGxR4(config-router)#neighbor 197.x.4.1 remote-as x
WGxR4(config-router)#neighbor 197.x.4.1 update-source loopback 0
WGxR4(config-router)#end
```

Answers to Review Questions:

Q1) Check the BGP table on router "Client." How many prefixes coming from your AS are in that BGP table?

25

Q2) Is there any other way of discovering how many prefixes you have advertised to the router "Client"?

show ip bgp neighbor *ip-address* **advertised**

Laboratory Exercise 5-2 Configure Filters in the BGP Transit Autonomous System

Use the following command sequence on your router WGxR4:

```
WGxR4(config-router)#neighbor 192.168.21.99 filter-list 1 in
WGxR4(config-router)#neighbor 192.168.21.99 filter-list 2 out
WGxR4(config)#ip as-path access-list 1 permit ^99$
WGxR4(config)#ip as-path access-list 2 permit ^$
WGxR4(config)#ip as-path access-list 2 permit _37$
WG1R4(config-router)#end
 WG1R4#clear ip bgp 192.168.21.99
```

Answers to Review Questions:

Q1) Why did you have to disable synchronization?

You should use BGP synchronization only in networks where BGP routes are redistributed in IGP. This design is no longer viable for service provider networks due to the large number of routes advertised in the Internet.

Q2) Why did you have to establish a full mesh of IBGP sessions?

IBGP split-horizon rules require a full mesh of IBGP sessions within the AS.

 Copyright © 2003, Cisco Systems, Inc.

Module 6: Scaling Service Provider Networks

Laboratory Exercise 6-1: BGP Route Reflectors

Use the following command sequence on your router WGxR1:

```
WGxR1(config)#router bgp x
WGxR1(config-router)#no neighbor 197.x.4.1 remote-as x
WGxR1(config-router)#no neighbor 197.x.6.1 remote-as x
WGxR1(config-router)#end
WGxR1#
```

Use the following command sequence on your router WGxR2:

```
WGxR2(config)#router bgp x
WGxR2(config-router)#bgp cluster-id 2 bgp cluster-id 102
WGxR2(config-router)#no neighbor 197.x.6.1 remote-as x
WGxR2(config-router)#neighbor 197.x.1.1 route-reflector-client
WGxR2(config-router)#end
```

Use the following command sequence on your router WGxR3:

```
WGxR3(config)#router bgp x
WGxR3(config-router)#bgp cluster-id 1 bgp cluster-id 101
WGxR3(config-router)#no neighbor 197.x.1.1 remote-as x
WGxR3(config-router)#neighbor 197.x.2.1 route-reflector-client
WGxR3(config-router)#end
```

Use the following command sequence on your router WGxR4:

```
WGxR4(config)#router bgp x
WGxR4(config-router)#no neighbor 197.x.2.1 remote-as x
WGxR4(config-router)#no neighbor 197.x.1.1 remote-as x
WGxR4(config-router)#end
```

Answers to Review Questions:

Q1) Did this design require you to configure a cluster-ID?

This design did not require specified cluster-IDs, because there are no redundant route reflectors in the network.

Q2) What is the default cluster-ID?

The default cluster-ID is the BGP router-ID.

Q3) When do you have to configure a cluster-ID?

You need to configure a cluster-ID in redundant route reflector designs.

Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise 6-2: BGP Confederations

Use the following command sequence on your router WGxR1:

```
WGxR1(config)#no router bgp x
WGxR1(config)#router bgp 65001
WGxR1(config-router)#network 192.168.x.0
WGxR1(config-router)#network 197.x.1.0
WGxR1(config-router)#network 197.x.8.0
WGxR1(config-router)#bgp confederation identifier x
WGxR1(config-router)#bgp confederation peers 65002
WGxR1(config-router)#neighbor 192.168.20.20 remote-as 20
WGxR1(config-router)#neighbor 192.168.20.22 remote-as 22
WGxR1(config-router)#neighbor 197.x.2.1 remote-as 65002
WGxR1(config-router)#neighbor 197.x.2.1 update-source loopback0
WGxR1(config-router)#neighbor 197.x.2.1 ebgp-multihop
WGxR1(config-router)#end
```

Use the following command sequence on your router WGxR2:

```
WGxR2(config)#no router bgp x
WGxR2(config)#router bgp 65002
WGxR2(config-router)#no synchronization
WGxR2(config-router)#network 197.x.2.0
WGxR3(config-router)#network 197.x.3.0
WGxR2(config-router)#bgp confederation identifier x
WGxR2(config-router)#bgp confederation peers 65001
WGxR2(config-router)#neighbor 197.x.1.1 remote-as 65001
WGxR2(config-router)#neighbor 197.x.1.1 update-source loopback0
WGxR2(config-router)#neighbor 197.x.1.1 ebgp-multihop
WGxR2(config-router)#neighbor 197.x.4.1 remote-as 65002
WGxR2(config-router)#neighbor 197.x.4.1 update-source loopback0
WGxR2(config-router)#end
```

Use the following command sequence on your router WGxR3:

```
WGxR3(config)#no router bgp x
WGxR3(config)#router bgp 65002
WGxR3(config-router)#no synchronization
WGxR3(config-router)#network 197.x.4.0
WGxR3(config-router)#network 197.x.5.0
WGxR3(config-router)#bgp confederation identifier x
WGxR3(config-router)#bgp confederation peers 65003
WGxR3(config-router)#neighbor 197.x.2.1 remote-as 65002
WGxR3(config-router)#neighbor 197.x.2.1 update-source loopback0
WGxR3(config-router)#neighbor 197.x.6.1 remote-as 65003
WGxR3(config-router)#neighbor 197.x.6.1 update-source loopback0
WGxR3(config-router)#neighbor 197.x.6.1 ebgp-multihop
WGxR3(config-router)#end
```

 Copyright © 2003, Cisco Systems, Inc.

Use the following command sequence on your router WGxR4:

```
WGxR4(config)#no router bgp x
WGxR4(config)#ip as-path access-list 10 permit ^99$
WGxR4(config)#ip as-path access-list 11 permit ^$|\)$|_37$
WGxR4(config)#router bgp 65003
WGxR4(config-router)#network 192.168.x.0
WGxR4(config-router)#network 197.x.6.0
WGxR4(config-router)#network 197.x.7.0
WGxR4(config-router)#bgp confederation identifier x
WGxR4(config-router)#bgp confederation peers 65002
WGxR4(config-router)#neighbor 192.168.21.99 remote-as 99
WGxR4(config-router)#neighbor 192.168.21.99 filter-list 10 in
WGxR4(config-router)#neighbor 192.168.21.99 filter-list 11 out
WGxR4(config-router)#neighbor 197.x.4.1 remote-as 65002
WGxR4(config-router)#neighbor 197.x.4.1 update-source loopback0
WGxR4(config-router)#neighbor 197.x.4.1 ebgp-multihop
WGxR4(config-router)#end
```

Answers to Review Questions:

Q1) What additional command did you have to use to establish intraconfederation EBGP sessions between loopback interfaces?

To establish intraconfederation EBGP sessions between loopback interfaces, you must specify ebgp-multihop on the intraconfederation EBGP neighbor.

Q2) Why was it necessary to change the AS-path filters?

The AS path of the local network has changed, because it includes the member-AS numbers.

Copyright © 2003, Cisco Systems, Inc.

 Copyright © 2003, Cisco Systems, Inc.

Module 7: Optimizing BGP Scalability

Laboratory Exercise 7-1: Limiting the Number of Prefixes Received from a BGP Neighbor

Use the following command sequence on your router WGxR1:

```
WGxR1(config)#router bgp 65001
WG1R1(config-router)#neighbor 192.168.20.22 maximum-prefix 15
warning-only
WG1R1#clear ip bgp 192.168.20.22

WG1R1(config)#router bgp 65001
WG1R1(config-router)#no neighbor 192.168.20.22 maximum-prefix 15
warning-only
WG1R1(config-router)#neighbor 192.168.20.22 maximum-prefix 15
WG1R1#clear ip bgp 192.168.20.22
```

Answers to Review Questions:

Q1) What can you configure on your router to allow the BGP session to restart automatically if the received number of prefixes exceeds the configured maximum?

restart parameter of the **maximum-prefix** command

neighbor *ip-address* maximum-prefix *maximum* restart

Laboratory Exercise 7-2: BGP Peer Groups

Use the following command sequence on your router WGxR1:

```
WG1R1(config)#router bgp 65001
WG1R1(config-router)#neighbor SvcPro peer-group
WG1R1(config-router)#neighbor 192.168.20.20 peer-group SvcPro
WG1R1(config-router)#neighbor 192.168.20.22 peer-group SvcPro
WG1R1(config-router)#neighbor SvcPro maximum-prefix 50
WG1R1(config-router)#neighbor SvcPro soft-reconfiguration in
WG1R1(config-router)#end
```

Answers to Review Questions:

Q1) If the number of prefixes received from both providers exceeds 50, what single command can you used to reset the BGP sessions on both routers?

clear ip bgp peer-group *peer-group name* **soft in**

In this example it would be:

clear ip bgp peer-group *SvcPro* **soft in**

Copyright © 2003, Cisco Systems, Inc.

Laboratory Exercise 7-3: BGP Route Dampening

Use the following command sequence on your router WGxR4:

```
WGxR4(config)#ip community-list 10 permit x:300
WGxR4(config)#ip prefix-list RFD24 permit 0.0.0.0/0 ge 24
WGxR4(config)#ip prefix-list RFD8 permit 0.0.0.0/0 ge 9
WGxR4(config)#route-map RFD permit 10
WGxR4(config-route-map)#match community 10
WGxR4(config-route-map)#route-map RFD permit 20
WGxR4(config-route-map)#match ip address prefix-list RFD24
WGxR4(config-route-map)#set dampening 20 750 2000 80
WGxR4(config-route-map)#route-map RFD permit 30
WGxR4(config-route-map)#match ip address prefix-list RFD8
WGxR4(config-route-map)#set dampening 15 750 2000 60
WGxR4(config-route-map)#route-map RFD permit 40
WGxR4(config-route-map)#set dampening 10 750 2000 40
WGxR4(config-route-map)#exit
WGxR4(config)#router bgp 65003
WGxR4(config-router)#bgp dampening route-map RFD
WGxR4(config-router)#end
```

Answers to Review Questions:

Q1) What is the purpose of route dampening?

Route dampening minimizes the impact of route flaps in downstream autonomous systems upon local and upstream autonomous systems.

Q2) Which routes are affected by route dampening?

Route dampening affects only EBGP routes.

 Copyright © 2003, Cisco Systems, Inc.

B

Course Glossary

The Course Glossary for *Configuring BGP on Cisco Routers* (BGP) v3.0 highlights and defines key terms and acronyms used throughout this course. Many of these terms are also described in the Cisco *Dictionary of Internetworking Terms and Acronyms* (ITA), available via http://www.cisco.com/univercd/cc/td/doc/cisintwk/ita/index.htm.

Acronym or Term	Expansion of Acronym	Short Definition for Mouseover	Definition of Acronym or Term	Source for Definition	Additional Explanation/ Description for this Course Context
AD	administrative distance	Administrative distance.	Rating of the trustworthiness of a routing information source. Administrative distance is often expressed as a numerical value between 0 and 255. The higher the value, the lower the trustworthiness rating.	ITA/Jan 2001	
aggregator		The last AS number that formed the aggregate route, followed by the IP address of the BGP speaker that formed the aggregate route.	An optional transitive attribute of length 6. The attribute contains the last AS number that formed the aggregate route, followed by the IP address of the BGP speaker that formed the aggregate route.	IETF[1] RFC 1771	
AS	autonomous system	A collection of networks under common administration.	A collection of networks under a common administration sharing a common routing strategy.	ITA/Jan 2001	
AS-path		A sequence of AS-path segments.	AS-path is a mandatory well-known attribute that is composed of a sequence of AS-path segments. This attribute identifies the autonomous systems through which routing information carried in this UPDATE message has passed.	IETF RFC 1771	
atomic aggregate		Used by a BGP speaker to inform other BGP speakers that the local system selected a less specific route without selecting a more specific route that is included in it.	A discretionary well-known attribute of length 0. It is used by a BGP speaker to inform other BGP speakers that the local system selected a less specific route without selecting a more specific route that is included in it.	IETF RFC 1771	

 Copyright © 2003, Cisco Systems, Inc.

Acronym or Term	Expansion of Acronym	Short Definition for Mouseover	Definition of Acronym or Term	Source for Definition	Additional Explanation/ Description for this Course Context
BGP	Border Gateway Protocol	The predominant interdomain routing protocol used on the Internet.	An interdomain routing protocol that replaces EGP. BGP4 is defined in RFC 1771. BGP4, the current version of BGP, is the predominant interdomain routing protocol used on the Internet.	ITA/Jan 2001	
black hole		An area of the internetwork where packets enter, but do not emerge.	Routing term for an area of the internetwork where packets enter, but do not emerge, due to adverse conditions or poor system configuration within a portion of the network.	ITA/Jan 2001	
CIDR	classless interdomain Routing	CIDR allows routers to group routes together to reduce the quantity of routing information carried by the core routers.	Technique supported by BGP4 and based on route aggregation. CIDR allows routers to group routes together to reduce the quantity of routing information carried by the core routers. With CIDR, several IP networks appear to networks outside the group as a single, larger entity. With CIDR, IP addresses and their subnet masks are written as four octets, separated by periods, followed by a forward slash and a two-digit number that represents the subnet mask.	ITA/Jan 2001	
client (route reflector)		Clients can have any number of EBGP sessions but may have only one IBGP session, the session with their route reflector.	Route reflector clients behave like traditional IBGP routers, but are excluded from the IBGP full mesh. Clients can have any number of EBGP sessions but may have only one IBGP session, the session with their route reflector.	"BGP Case Studies Section 4" http://www.cisco.com/warp/public/459/bgp-toc.html - routereflectors	

Acronym or Term	Expansion of Acronym	Short Definition for Mouseover	Definition of Acronym or Term	Source for Definition	Additional Explanation/ Description for this Course Context
cluster (route reflector)		A group of redundant route reflectors and their clients.	A group of redundant route reflectors and their clients form a route reflector cluster.	"BGP Case Studies Section 4" http://www.cisco.com/warp/public/459/bgp-toc.html - routereflectors	
community		A group of destinations that share a common set of properties to which routing decisions can be applied.	In the context of BGP, a community is a group of destinations that share a common set of properties to which routing decisions (such as acceptance, preference, and redistribution) can be applied.	http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm#xtocid10	
community internet		The route to which this community is applied should be advertised to the Internet community.	A predefined BGP community attribute indicating that the route to which this community is applied should be advertised to the Internet community; all routers belong to it.	http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm#xtocid10	
community no-advertise		The route to which this community is applied should not be advertised to any peer.	A predefined BGP community attribute indicating that the route to which this community is applied should not be advertised to any peer.	http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm#xtocid10	
community no-export		The route to which this community is applied should not be advertised to EBGP peers.	A predefined BGP community attribute indicating that the route to which this community is applied should not be advertised to EBGP peers.	http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm#xtocid10	
confederation		Grouping of multiple subautonomous systems.	A BGP technique used to reduce the IBGP full mesh within an autonomous system by dividing an autonomous system into multiple subautonomous systems and grouping them into a single confederation.	Cisco IOS® 12.2 documentation: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2//1cfbgp.htm - xtocid44	

 Copyright © 2003, Cisco Systems, Inc.

Acronym or Term	Expansion of Acronym	Short Definition for Mouseover	Definition of Acronym or Term	Source for Definition	Additional Explanation/ Description for this Course Context
convergence		Agreement on the topology of an internetwork after a change in that topology.	Speed and ability of a group of internetworking devices running a specific routing protocol to agree on the topology of an internetwork after a change in that topology.	ITA/Jan 2001	
dampening		A mechanism to minimize the instability caused by route flapping and oscillation over the network.	A mechanism to minimize the instability caused by route flapping and oscillation over the network.	http://www.cisco.com/warp/public/459/16.html#A24.4	
distance vector		Class of routing algorithms that iterate on the number of hops.	Class of routing algorithms that iterate on the number of hops in a route to find a shortest-path spanning tree. Distance vector routing algorithms call for each router to send its entire routing table in each update, but only to its neighbors.	ITA/Jan 2001	
DMZ	demilitarized-zone	Outside the firewall, but is still available for use by the internal (protected) network.	DMZ is a term first used in complex, multiple-machine firewall setups, where a computer is placed outside the firewall, but is still available for use by the internal (protected) network.	http://www.dslreports.com/information/kb/DMZ	
EBGP	External BGP	BGP used to exchange routing information between different autonomous systems.	When BGP is used to exchange routing information between different autonomous systems, the protocol is referred to as External BGP (EBGP).	http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm#xtocid11	
EGP	exterior gateway protocol	Any routing protocol used to exchange information between autonomous systems.	Any Internet protocol used for exchanging information between autonomous systems. Examples: BGP, EGP	ITA/Jan 2001	

Acronym or Term	Expansion of Acronym	Short Definition for Mouseover	Definition of Acronym or Term	Source for Definition	Additional Explanation/ Description for this Course Context
flap, flapping		Problem caused by intermittent interface failures.	Routing problem where an advertised route between two nodes alternates (flaps) back and forth between two paths due to a network problem that causes intermittent interface failures.	ITA/Jan 2001	
IBGP	Internal BGP	BGP when used to exchange routing information within an autonomous system.	When BGP is used to exchange routing information within an autonomous system, the protocol is referred to as Internal BGP (IBGP).	http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm#xtocid11	
IGP	Interior Gateway Protocol	Any routing protocol used to exchange routing information within an autonomous system.	Any Internet protocol used to exchange routing information within an autonomous system. Examples: Routing Information Protocol (RIP), Open Shortest Path First (OSPF), Intermediate System-to-Intermediate System (IS-IS)	ITA/Jan 2001	
interdomain		Between two or more logical domains.	Term used to describe routing between two or more logical domains.	ITA/Jan 2001	
ISP	Internet service provider	Internet service provider.	Company that provides Internet access to other companies and individuals.	ITA/Jan 2001	
local preference		Used to inform other BGP speakers within an AS of the degree of preference of the originating speaker for an advertised route.	A discretionary well-known attribute that is a four–octet, nonnegative integer. It is used by a BGP speaker to inform other BGP speakers in its own autonomous system of the degree of preference of the originating speaker for an advertised route.	IETF RFC 1771	
MD5	Message Digest 5	One-way hashing algorithm.	A one-way hashing algorithm that produces a 128-bit hash.	ITA/Jan 2001	

 Copyright © 2003, Cisco Systems, Inc.

Acronym or Term	Expansion of Acronym	Short Definition for Mouseover	Definition of Acronym or Term	Source for Definition	Additional Explanation/ Description for this Course Context
MED	Multi Exit Discriminator	Used to discriminate among multiple exit points to a neighboring autonomous system.	This is a nontransitive optional attribute that is a four–octet, nonnegative integer. The value of this attribute may be used by the decision process of a BGP speaker to discriminate among multiple exit points to a neighboring autonomous system.	IETF RFC 1771	
member-AS (sub-AS)	member autonomous system	Autonomous system internal to a BGP confederation.	An autonomous system internal to a BGP confederation.		
MTU	maximum transmission unit	Maximum packet size, in bytes.	Maximum packet size, in bytes, that a particular interface can handle.	ITA/Jan 2001	
multihomed AS	multihomed autonomous system	A single AS that is attached to multiple autonomous systems.	A single autonomous system that is attached to multiple autonomous systems.		
NAT	Network Address Translation	Network Address Translation.	Mechanism for reducing the need for globally unique IP addresses. NAT allows an organization with addresses that are not globally unique to connect to the Internet by translating those addresses into globally routable address space. Also known as “Network Address Translator.”	ITA/Jan 2001	
next-hop		The IP address of the border router that should be used as the next hop to the destinations listed.	This is a mandatory well-known attribute that defines the IP address of the border router that should be used as the next hop to the destinations listed in the Network Layer Reachability field of the UPDATE message.	IETF RFC 1771	
nonclients (route reflector)		IBGP peers of a route reflector that are not configured as clients.	Other IBGP peers of the route reflector that are not configured as clients.	BGP Case Studies http://www.cisco.com/warp/public/459/bgp-toc.html - routereflectors	

Acronym or Term	Expansion of Acronym	Short Definition for Mouseover	Definition of Acronym or Term	Source for Definition	Additional Explanation/ Description for this Course Context
origin		The origin of the path information.	A mandatory well-known attribute that defines the origin of the path information. The origin attribute can be set to “IGP,” “EGP,” or “Incomplete.”	IETF RFC 1771	
POP	point of presence	Point of presence.	A physical location where an interexchange carrier installs equipment to interconnect with a local exchange carrier (LEC).	ITA/Jan 2001	
route reflector		A router that advertises (reflects) IBGP-learned routes to other IBGP speakers.	A BGP technique used to reduce the IBGP full mesh within an autonomous system by modifying split-horizon rules to allow a router to advertise (reflect) IBGP-learned routes to other IBGP speakers.	BGP Case Studies Section 4 http://www.cisco.com/warp/public/459/bgp-toc.html - routereflectors	
router-ID	router identifier	Highest IP address on the router or the highest loopback interface.	The router-ID is set to the IP address of a loopback interface if one is configured. If no virtual interfaces are configured, the highest IP address is configured for a physical interface on that router. Peering sessions will be reset if the router-ID is changed.	Cisco IOS 12.2 documentation	
subnet		Arbitrary segment of a network to provide a multilevel, hierarchical routing structure.	In IP networks, subnetworks are networks arbitrarily segmented by a network administrator in order to provide a multilevel, hierarchical routing structure while shielding the subnetwork from the addressing complexity of attached networks.	ITA/Jan 2001	
supernet		Aggregation of IP network addresses advertised as a single classless network address.	Aggregation of IP network addresses advertised as a single classless network address.	ITA/Jan 2001	

 Copyright © 2003, Cisco Systems, Inc.

Acronym or Term	Expansion of Acronym	Short Definition for Mouseover	Definition of Acronym or Term	Source for Definition	Additional Explanation/ Description for this Course Context
transit AS	transit autonomous system	A multihomed AS that can be used for transit traffic.	A multihomed autonomous system that can be used for transit traffic, that is, traffic with an origin and destination that do not belong to the AS.	*Internet Routing Architectures*, Halabi. Cisco Press	
weight (metric)		Preference assigned locally to the router and not propagated or carried through any BGP route updates.	Weight is a Cisco-defined BGP attribute used for a best-path selection process. The weight is assigned locally to the router and is not propagated or carried through any BGP route updates. Weight can be a number from 0 to 65535. Paths that the router originates have a weight of 32768 by default, and other paths have a weight of 0.	"BGP Case Studies Section 2" (Cisco documentation)	

[1]IETF = Internet Engineering Task Force

 Copyright © 2003, Cisco Systems, Inc.